DIPLOMARBEIT

Titel der Diplomarbeit

Differential Equation Models for Surface Reactions of SnO_2 Nanowire Gas Sensors and their Inverse Modeling

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Verfasserin:	Marina Rehrl
Studienkennzahl lt. Studienblatt:	A 405
Studienrichtung lt. Studienblatt:	Mathematik
Betreuer:	Univ.-Ass. Priv.-Doz. DI Dr. Clemens Heitzinger

Wien, August 2011

Overview

In the line of the Vienna Science and Technology Fund (WWTF) project "Mathematics and Nano-Sensors" a working relationship between the Austrian Institute of Technology (AIT) and the Wolfgang Pauli Institute (WPI) was established to build and test AIT fashioned tin oxide nanowire gas sensors and provide mathematical models to predict the outcome of gas measurements. As state of the art gas sensors still suffer from low selectivity, the detailed modeling of the surface reactions to different test gases is essential to overcome this issue.

Despite the considerable amount of papers giving detailed discussions of experimental results in various gas atmospheres (a critic review can be found in [Comini et al. 2009]), the discussed surface reaction models in the literature mainly concentrate on more simple test gases like oxygen, carbon oxide and inert gases. Therefore we were highly interested in providing surface models for additional gases like, for example hydrogen sulfide or nitrogen dioxide, which are known to be important for environmental control systems and various other applications.

The developed response models, described in this work, are composed of a surface reaction model and a charge transport model and predict the change of conductance of the sensor upon changes in the thermal and chemical environment. Different transport models in the form of functional equations were derived for different types of sensor structures. The surface reaction models consist of rate equations in the form of a system of parameter dependent ordinary differential equations (ODEs), which differ depending on temperature, type of modeled gas as well as possible interactions with other gases.

In order to facilitate the simulation of the sensor response, the theory of inverse modeling of dynamic models, with special regard to the estimation of model parameters through nonlinear least squares estimators and suitable optimization algorithms, was discussed.

The simulation of the response of a gas sensor with a bundle of a nanowires as sensing element in an inert atmosphere was accomplished by the fitting of 1 parameter in the charge transport model and 5 parameters in the surface state model through nonlinear least squares estimation.

This work is therefore organized as follows:

Chapter 01 intends to familiarize the reader with the technical aspect of gas sensor and gives an overview to their functionality.

Chapter 02 derives charge transport models for currently researched sensor architectures and discusses their properties and applicability. Depending on the state of carrier depletion in the sensor element different transport models were considered.

Chapter 03 gives a detailed derivation and discussion of parameter dependent ordinary differential equation models that describe the chemical reactions between different types of gases and the surface of a tin oxide sensor. Detailed surface reaction models were derived for the most important test gas species.

Chapter 04 discusses the different types of surface reaction models and structures them into a hierarchy starting with intrinsic model, oxygen model to combinations of test gas models. Then we will analyze the qualitative properties of the most important surface models to prove the existence and uniqueness of their solution.

Chapter 05 addresses the estimation of parameters for models consisting of ordinary differential equations. The standard form of an dynamical model is stated to aid the discussion of the general course of action for estimating an ODE model with regard to a suitable optimization algorithm. In the second part we will cover the theory of Nonlinear Least Squares Estimators and discuss and prove their asymptotic properties.

Chapter 06 presents the quantitative analysis of the derived charge transport models in combination with the intrinsic surface state trapping model. The response of a sensor consisting of a bundle of nanowires in an inert atmosphere was simulated and the parameters for the surface reaction model as well as the charge transport model were fitted.

Contents

Chapter 1

Introduction

Gas sensors are a type of chemical sensors, which are devices able to convert chemical states into electrical signals, and are used to detect the concentration and hereby the presence of certain target gases in the atmosphere. The detection of gases is important for many fields of applications:

- safety engineering,
- health care,
- bio-sciences and
- environmental monitoring

if only to name a few. A quiet extensive (but still not exhaustive) list of current and possible future applications can be found in [Zima 2009].

To meet this demand for the applicability to different fields, considerable research into new types of sensors is needed, including efforts to enhance the performance and understand the working principles of these sensor devices. Necessary properties of gas sensor are selectivity (i.e., response only to the targeted gas) and sensitivity (i.e., providing of sufficiently measurable sensor response). Current gas sensors however are highly cross sensitive, especially regarding gas species that show similar reducing or oxidizing properties. The mathematical modeling of the underlying sensing mechanisms, which is the modeling of the occurring surface reactions, helps in the scientific understanding of these devices and is the key for further advancement of the sensors to overcome the issue of cross selectivity.

The principle of gas detection with SnO_2, or nearly every other metal oxide semiconductor, is based on the measurable change of the electrical conductance upon the adsorption of gas species, as there usually occurs a charge transfer between adsorbed gas molecules and the sensor surface. As gas sensors are usually operated in an oxygen atmosphere, adsorbed oxygen extracts electrons from the semiconductor and, if the sensing material conducts by electrons, decreases the conductance. If, for example, carbonic target gas species are present in the

atmosphere, they become ionized by pre-adsorbed negatively charged oxygen, when reacting with the sensor surface, and return electrons into the solid. This leads to an increase in conductivity. The conductance of a sensor is therefore much higher if a type of reducing gas, like carbon monoxide, is present in the atmosphere then it is in "clean" air. Because of the large surface-to-volume ratio of a nanowire-structured sensor device, the response of a nanowire gas sensor is also very sensitive to changes in chemistry and dielectric properties of the surface, caused by this exchange of charge carriers.

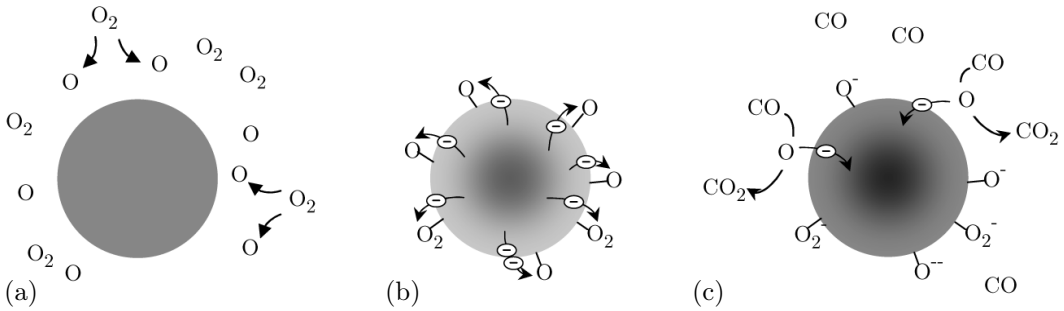


Figure 1.1: A schematic representation of the interaction of a SnO_2 nanowire with O_2 and CO. Figure (a) shows oxygen, in atomic and molecular form in the atmosphere around the nanowire. In the second figure oxygen species are bound to the surface of the nanowire by extracting electrons from the bulk material. Figure (c) shows the effect of CO as it extracts the oxygen species from the surface and reintroduces the previously bound electrons into the nanowire.

It was known for many years that the electrical properties of semiconductors are sensitive to ambient gases. After physical studies done by Brattain&Bardeen and Morrison in 1953 and since Seiyama et al. discovered in 1962 that the presence of reactive gas species in the atmosphere causes a tremendous change in the electrical conductivity of ZnO, the use of metal oxides semiconductors for gas sensing purposes was intensively studied.

Tin oxide arose as an especially favorable material for these purposes as it possesses a high sensitivity to various target gases and is a generally well understood and easily fabricated material [Zima 2009]. SnO_2 was also used for the first commercial gas sensors, a Taguchi-type sensor, manufactured by Figaro in Japan.

The semiconducting behavior of tin oxide arises, as is typical for metal oxides, from deviations of the stoichiometry of the material, as otherwise ideal stoichiometric SnO_2 would be an isolator at room temperature. Therefore the termination of the periodic structure at the surface of tin oxide may form surface-localized electronic states within the semiconductor band gab. The n-conducting properties of the tin oxide stems from the electron donor effect of the oxygen surface defects (e.g., oxygen vacancies) in the crystal lattice, which can be singly and doubly ionized and play a significant role in the process of gas sensing. Those oxygen vacancies are mainly formed in the manufacturing process, as oxygen atoms may escape into

the atmosphere. The appearance of these acceptor states at the surface induce charge transfer between the SnO_2 bulk material and the surface to establish thermal equilibrium. This charge transfer results in a region, depleted of charge carriers (i.e., electrons), under the surface of the material, called the surface space charge region, which reaches approximately a few Debye length into the bulk material [Comini et al. 2009]. Upon exposure to atmospheric oxygen, more charge carriers are trapped into energy levels at the surface and the depletion region is amplified.

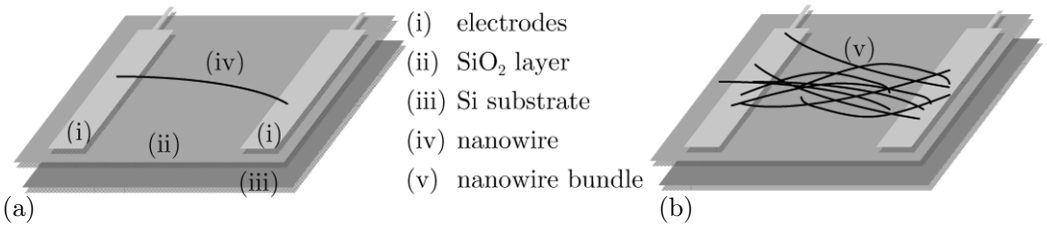


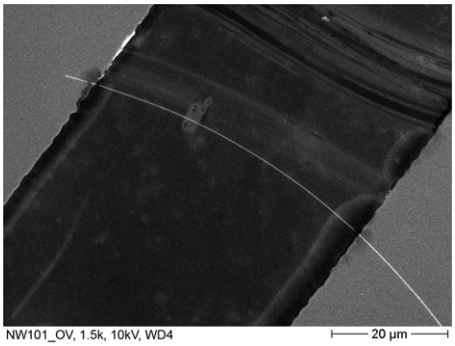
Figure 1.2: A schematic representation of the architecture of a nanowire sensor consisting of a single nanowire (a) and bundle or network of nanowires (b).

Is the gas sensor then of nano-sized dimensions (i.e., comparable in size to the Debye length), the whole device is nearly depleted of charge carriers and exhibits a much poorer conductivity than micro-sized sensors in ambient air. Thus, when the nano-sized sensor is exposed to a target gas, the resulting conductance change is much greater, as electrons released from surface states have comparably a much greater impact on the conductance than in micro-sized sensors [Huang and Wan 2009].

The architecture of such a nanowire gas sensor comprises of a single or a bundle of nanowires bridging two metal electrodes (contacts) on a silicon substrate covered with SiO_2, acting as an insulating layer between the nanowire/electrode structure and the conducting silicon, as can be seen in Figure 1.2.

The diameter of a nanowire lies in the scale of the gas species we wish to detect. Therefore every single adsorbed gas molecule leads to a measurable change in the sensor conductance, especially when considering the high surface-to-volume ratio of nanowires. A high sensitivity to comparably small gas concentrations as well as a the small size are definite advantages of nanowire gas sensors. Furthermore small response and recovery times as well as low production costs are prominent advantages of nano-sized gas sensors [Tischner et al. 2009].

Although nanowire sensor devices are a definite advancement compared to traditional nano-scale thin film sensors [Tischner et al. 2008; Zima et al. 2010], they still face some of the common problems of those types of sensors. Noteworthy are their somehow unpredictable response to the presence of humidity, their poor longtime stability and reproducibility and the already mentioned low selectivity.

(a) 

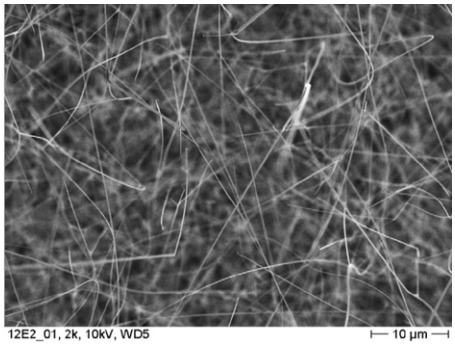
(b) 

Figure 1.3: Picture of a single nanowire (a) and a bundle/network of nanowires (b). The single nanowire is positioned between two contacts. According to References [Köck et al. 2009] and [Tischner et al. 2009].

For definitions to chemical and physical terminologies and concepts in this and the following chapters, the interested reader is referred to the work of [Zima 2009].

Chapter 2

Charge Transport in the Sensors

The mechanism of charge transfer in a semiconductor gas sensor depends strongly on its morphology. If the sensitive part of a sensor structure is fashioned as a sensor film, a general distinction between compact and porous layers can be made. Compact sensing layers exclude the possibility of gas penetrating the sensing layer. In this case the interaction with gas takes place only on the geometric surface, where else porous sensing layers allow for the gas to access the sensing layer. Their active surface reaches therefore into the sensing layer.

For compact sensing layers or sensors consisting of single nanowires a formula for the conductance, as the product of conductivity and a geometry factor, for electrical devices with a uniform cross section is applicable. When combined with an additional equation for the density of electrons on the sensor surface the resulting conductance formula is applicable to sensor with grain boundaries. This approach is referred to as the Potential Barrier Theory (PBT) by a number of authors, [Ding et al. 2001; Morrison 1990].

The Thermoelectronic Emission Theory (TEET), also called Diode Theory, as well as the Diffusion Theory (DT), which can be found in [Bârsan and Weimar 2001; Broniatowski 1985], explain the link between surface phenomena and the measured conductance for porous sensor films with grains larger than the Debye length λ_D.

All the above theories will then lead to an expression of the conductance as a function of the surface potential barrier V_s. A further equation is therefore needed to link the value of V_s to the concentration of surface states N_s, which are derived from surface reaction models in Chapter 3.

Actually, both sensor constructions can be divided into two additional cases — thick and thin sensing layers/nanowires, where the deciding factor is the ratio between the layer/wire thickness and the Debye length. Depending on the actual morphology of the grain-grain contact region the respective transport mechanism is determined. But ultimately the resulting conductance formulas for the above mentioned sensor structures are basically the same and only differ in a multiplicative constant.

2.1 Potential Barrier Theory

The existence of surface states at a surface gives rise to a difference between the energy level of the surface state and the conduction band of the bulk material. This enables electrons to alternate between conduction band and surface states. The conduction electrons, which were able to cross over and thereafter occupy a surface state, create a repulsive potential barrier at the surface to further prevent the trapping of conduction electrons into unoccupied surface states. Such potential barriers are a particular hindrance for electrons when they occur at the contact areas between single grains, which are anyway rather narrow. In order to contribute to the conductance, these electrons have to acquire enough kinetic energy to overcome the surface potential barrier V_s while traveling through the grain contact areas.

When discussing the conductance G, one has to start with the microscopic conductivity σ. The electric conductivity of a semiconductor crystal has a electronic/hole and a ionic part. However, as tin oxide gas sensors are usually operated in temperatures between 200° and 400°C, the ionic component of the conductivity can be neglected [Zima 2009]. Furthermore, because of SnO_2 being a n-type semiconductor, one refers to the electronic part of the overall conductivity/conductance. The electronic conductivity in a homogeneous single crystal is therefore given by

$$\sigma = \sigma_e + \sigma_p + \sum \sigma_{ion,i} \approx \sigma_e + \sigma_p \approx \sigma_e = q \cdot \mu \cdot n,$$

where q gives the elementary charge, n is defined as the electron concentration and μ states their mobility.

The relation between the conductivity σ and the conductance G for the case of an n-type semiconductor is given by

$$G = const \cdot \sigma = const \cdot q \cdot \mu \cdot n_s \tag{2.1}$$

for a constant factor *const* depending on the geometry of the sensor. As the conductance is dominated by the electron transfer across the surface potential barrier at the inter-granular contact regions, G is proportional to the density n_s of electrons that are responsible for the conduction at the surface. The density of electrons n can therefore be set to n_s

To complete this relation we have to derive a equation to set the density of surface electron in a relation to the potential barrier at the surface.

The Density of Electrons at the Surface n_s

To calculate the electron density n_s on the surface of a semiconducting material we have to start with the bulk value n_b. First we consider the density of charge carriers or occupied states per unit volume and energy $n(E)$, which is simply the product of the density of available states in the conduction band $g_c(E) = \frac{8\pi\sqrt{2}}{h^3} m_e^{\star\frac{3}{2}} \sqrt{E - E_F}$ and the probability that each of these states is occupied, which is the Fermi-Dirac probability function $f(E)$. The density of occupied states is therefore given by

$$n(E) = g_c(E) \cdot f(E).$$

The density of electrons in the semiconductor bulk material is then obtained by integrating the density of carriers over all possible energies within a band, as seen in equation

$$\begin{aligned} n_b &= \int_{\text{bottom of conduction band}}^{\text{top of conduction band}} g_c(E) \cdot f(E) dE \\ &= \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{h^3} m_e^{\star\frac{3}{2}} \sqrt{E - E_F} \cdot \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} dE. \end{aligned}$$

As the Fermi function converges to zero for rising energies the actual location of the top of the conduction band needs not be known and can be replaced by infinity. Assuming that the Fermi level is at least $2kT$ away from the conduction band edge (i.e., $E_c - E_F > 2kT$), this allows for the replacement of the Fermi distribution by the Boltzmann distribution. These considerations leads us to

$$\begin{aligned} n_b &= \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{h^3} m_e^{\star\frac{3}{2}} \sqrt{E - E_F} \cdot \exp\left(\frac{E_F - E}{kT}\right) dE \\ &= \frac{8\pi\sqrt{2}}{h^3} m_e^{\star\frac{3}{2}} \int_{E_c}^{\infty} \sqrt{E - E_F} \cdot \exp\left(\frac{E_F - E}{kT}\right) dE \\ &= \frac{8\pi\sqrt{2}}{h^3} m_e^{\star\frac{3}{2}} \cdot \frac{kT^{\frac{3}{2}}}{2} \sqrt{\pi} \cdot \exp\left(\frac{E_F - E_C}{kT}\right) \\ &= N_c \exp\left(\frac{E_F - E_C}{kT}\right) \end{aligned}$$

with $N_c = 2\left(\frac{2\pi m_e^\star kT}{h^2}\right)^{\frac{3}{2}}$ the effective density of states in the conduction band. The density of electrons at the surface of an n-type semiconductor n_s is then given by the previous equation multiplied by a Boltzmann factor, taking the potential barrier at the surface into account. Therefore, the density n_s of free electrons energized enough to overcome the surface potential

barrier V_s, is given by

$$n_s = N_c \exp\left(-\frac{(qV_s + E_c - E_F)}{kT}\right) = n_b \exp\left(-\frac{qV_s}{kT}\right), \tag{2.2}$$

according to [Morrison 1990]. The conduction formula according to the Potential Barrier Theory becomes therefore

$$G = const \cdot q\mu n_b \exp\left(-\frac{qV_s}{kT}\right). \tag{2.3}$$

2.2 Diffusion Theory

If the depletion layer is large compared to the mean free path of electrons, the Diffusion Theory is applicable, as in this case the concepts of drift and diffusion are valid. To derive a equation for the conductance we will consider the potential barriers on the two sides of the grain boundary separately and then join the solutions under the condition of current continuity. We consider the case of a n-type semiconductor like SnO_2. Then the Diffusion Theory gives the current density across the grain boundary as the sum of contributions from the dependence of the current on field intensity and diffusion resulting from the gradient in the carrier density:

$$\begin{aligned} j(t) &= \sigma(x,t)E(x,t) + qD\frac{\partial n}{\partial x} \\ &= -q\mu n(x,t)\frac{\partial V(x,t)}{\partial x} + \mu kT\frac{\partial n(x,t)}{\partial x} \end{aligned}$$

for

$D = \frac{\mu kT}{q}$ ..... carrier diffusion coefficient,
$V(x,t)$ ..... electrostatic potential,
$E(x,t) = -\frac{\partial V(x,t)}{\partial x}$ ..... electric field.

As solution of the above equation, the following expression for the current density in the diffusion case is given, [Broniatowski 1985], by

$$j(t) = q\mu E_i n_b \left(\exp\left(-\frac{q\,(V_a + V_i)}{kT}\right) - \exp\left(-\frac{qV_i}{kT}\right)\right)$$

for E_i the electric field strength on the i^{th} side. Equating the current density for the left side $i = 1$ and right side $i = 2$ gives us

$$E_1\left(\exp\left(-\frac{q\,(V_a + V_i)}{kT}\right) - \exp\left(-\frac{qV_1}{kT}\right)\right) = E_2\left(\exp\left(-\frac{q\,(V_a + V_i)}{kT}\right) - \exp\left(-\frac{qV_2}{kT}\right)\right),$$

$$E_2 \exp\left(-\frac{qV_2}{kT}\right) - E_1 \exp\left(-\frac{qV_1}{kT}\right) = (E_2 - E_1) \exp\left(-\frac{q\,(V_a + V_i)}{kT}\right),$$

which in turn leads to

$$\exp\left(-\frac{q\,(V_a + V_i)}{kT}\right) = \frac{E_2 \exp\left(-\frac{qV_2}{kT}\right) - E_1 \exp\left(-\frac{qV_1}{kT}\right)}{E_2 - E_1}.$$

Insertion of the last relation into the equation for charge density for one side of the barrier leads us to

$$\begin{aligned} j(t) &= \frac{q\mu E_2 n_b}{E_2 - E_1}\left(E_2 \exp\left(-\frac{qV_2}{kT}\right) - E_1 \exp\left(-\frac{qV_1}{kT}\right) - (E_2 - E_1) \exp\left(-\frac{qV_2}{kT}\right)\right) \\ &= \frac{q\mu E_1 E_2 n_b}{E_2 - E_1}\left(\exp\left(-\frac{qV_2}{kT}\right) - \exp\left(-\frac{qV_1}{kT}\right)\right) \\ &= \frac{q\mu E_s n_b}{2} \exp\left(-\frac{qV_2}{kT}\right)\left(1 - \exp\left(-\frac{qV_a}{kT}\right)\right), \end{aligned}$$

where we set the applied voltage $V_a = V_1 - V_2$ to zero to get the barrier height V_s as well as field strength $E_s = E_1 = -E_2$ at the boundary. When we differentiate the current density with respect to V_a we get the slope of the current-voltage characteristics, which is the conductivity (according to [Taylor et al. 1952]), when we set the applied voltage V_a to zero. The conductivity is therefore given by

$$\sigma = \left.\frac{dj}{V_a}\right|_{V_a=0} = \frac{q^2\mu}{2kT} E_s n_b \exp\left(\frac{-qV_s}{kT}\right) = \frac{q^{5/2} n_b^{3/2} \mu V_s^{1/2}}{\sqrt{2\varepsilon} kT} \exp\left(\frac{-qV_s}{kT}\right)$$

as E_s equals $\sqrt{\frac{2qn_bV_s}{\varepsilon}}$ according to [Broniatowski 1985]. By multiplying the conductivity with an factor *const*, which is the effective area seen by electrons while traveling from grain to grain, we obtain

$$G_{DT} = const\ \frac{q^{5/2} n_b^{3/2} \mu V_s^{1/2}}{\sqrt{2\varepsilon} kT} \exp\left(\frac{-qV_s}{kT}\right), \tag{2.7}$$

the conduction formula of the Diffusion Theory.

2.3 Thermoelectronic Emission Theory

For the case that the surface potential barrier width is much smaller than the mean free path of electrons ($\lambda \geq x_0$), the current density is proportional to the difference of the electron fluxes crossing the interface. Therefore only those electrons with kinetic energy equal or larger than

the potential barrier height are able to cross the boundary. This is reflected in the current density

$$j(t) = q n_b v_{th} \left(\exp\left(-\frac{qV_2}{kT}\right) - \exp\left(-\frac{qV_1}{kT}\right)\right),$$

where the mean thermal velocity of electrons v_{th} is given by $\sqrt{\frac{8kT}{\pi m^\star}}$ and $m^\star$ denotes the effective mass of electrons. As the applied voltage or bias is given by $V_a = V_1 - V_2$ we can rewrite the current density as

$$j(t) = q n_b v_{th} \exp\left(-\frac{qV_2}{kT}\right)\left(1 - \exp\left(-\frac{qV_a}{kT}\right)\right).$$

Differentiation of the current density with respect to V_a results in the electrical conductivity

$$\sigma = \left.\frac{dj}{V_a}\right|_{V_a=0} = \frac{q^2 n_b v_{th}}{kT} \exp\left(-\frac{qV_s}{kT}\right)$$

when we set the applied voltage to zero. By multiplication with an area factor we get the conduction formula for the Thermoelectronic Emission Theory

$$G_{TEET} = const \cdot \frac{q^2 n_b v_{th}}{kT} \exp\left(-\frac{qV_s}{kT}\right). \tag{2.8}$$

The previous section revealed that the conductance formulas derived for Potential Barrier Theory, Diffusion Theory as well as Thermoelectronic Emission Theory are of a similar form

$$G = const \cdot q\ v\ n_b\ \exp\left(-\frac{qV_s}{kT}\right), \tag{2.9}$$

while v denotes a velocity term given by

$$v = \begin{cases} \mu & \text{Potential Barrier Theory} \\ \frac{q}{2kT}\ \mu E_s & \text{Diffusion Theory} \\ \frac{q}{kT}\ v_{th} & \text{Thermoelectronic Emission Theory,} \end{cases}$$

whose exact form depends on the underlying theory.

In order to relate the above conductance formula to the density of occupied surface states N_s, which in the following chapter is obtained from surface states models, we need to derive the surface potential barrier V_s as a function of N_s. For this purpose we also have to consider

the state of charge depletion in the respective nanostructure.

2.4 The Surface Potential Barrier V_s and Depletion of Charge Carriers

The actual link between the measured quantity of conductance and the occupied surface states N_s is given by either the Schottky relation or a more close examination of the geometry of the sensor structure. The choice of the type of approach depends on the width of the nanostructure w in relation to the width of its depletion region, which scale can be described by the Debye length λ_D. The case $\lambda_D \ll w$ applies to sensors whose depletion region does not reach very far into the nanostructure. These sensing areas have therefore a inner region unaffected by surface phenomena. In the case of $\lambda_D \approx w$, the sensor is completely depleted of bulk electrons and surface phenomena therefore influence the conductance of the whole sensing area.

In both cases we will derive an expression for the surface potential barrier in terms of the surface state density, which will then be used to include the value of N_s into the already derived conductance formulas.

2.4.1 Partial Depletion

For nanostructures which are not completely depleted of charge carriers ($\lambda_D \ll w$), we derive the Schottky relation from the one-dimensional Poisson equation

$$\frac{d^2\psi}{dx^2} = \frac{qN_i}{\varepsilon},$$

with ψ the electrical potential and N_i the ion density in the space-charge region. The Poisson equation describes the change in the electrical potential as a function of the distance through the space charge region. We change the coordinates to $V(x) = \psi_0 - \psi(x)$, with ψ_0 the electrical potential in the bulk material, in order to gain the relation

$$V = \frac{qN_i\,(x - x_0)^2}{2\varepsilon},$$

after solving the Poisson equation with the boundary condition $\frac{dV}{dx} = V = 0$ for $x = x_0$, while x_0 denotes the thickness of the depletion layer, $0 \leq x \leq x_0$. As for n-type material $N_i = N_D$, which is the density of electrons in the semiconductor. The factor $N_D x_0$ denotes then the number of electrons extracted from the space charge region. As the electroneutrality condition states that the charge in the depletion layer equals the charge captured on the surface, [Bârsan and Weimar 2001], the product $N_D x_0$ also denotes the density of charged

surface states N_s. This calculation leads us finally to the Schottky relation

$$V_s = \frac{qN_s^2}{2\varepsilon N_D}, \tag{2.10}$$

with V_s the surface potential barrier for $x = 0$, where ε states the electrical permittivity of the semiconductor material.

According to [Madou and Morrison 1989] all the donors in SnO_2 are ionized at room temperature and above so the density of ionized donors N_D can be considered constant, [Ding et al. 2001; Fort et al. 2006b]. Then, according to [Bârsan and Weimar 2001], the density of electrons in the bulk for a semiconductor with completely ionized donors and little acceptors equals the density of donors:

$$n_b = N_D. \tag{2.11}$$

2.4.2 Complete Depletion

For thin nanostructures with a scale of diameter comparable to the Debye length ($\lambda_D \approx w$), the Schottky relation is not applicable. In this case we first consider the nanowire as a cylinder with radius R, a depletion region of thickness x_0 and charge density equal to N_D, as was proposed by [Fort et al. 2010]. We introduce cylindrical coordinates and write the electrical field in the depletion region as

$$E(r) = \frac{qN_D}{2\varepsilon}\left(r - \frac{(R-x_0)^2}{r}\right)$$

and get the electrical potential

$$V(r) = \frac{qN_D}{2\varepsilon}\left(\frac{r^2}{2} - (R-x_0)^2\ln(r)\right)$$

from the relationship $\frac{dV(r)}{dr} = E(r)$. This implies that the surface potential barrier V_s is the difference between the electrical potential on the surface and the potential at the begin of the depletion region

$$\begin{aligned} V_s &= V(R) - V(R-x_0) \\ &= \frac{qN_D}{2\varepsilon}\left(\frac{2Rx_0 - x_0^2}{2} - (R-x_0)^2\ln\left(\frac{R}{R-x_0}\right)\right) \\ &= \frac{qN_D}{4\varepsilon}R^2 \qquad \text{for } R \approx x_0. \end{aligned} \tag{2.12}$$

So, in the case of a complete depletion of the wire ($R \approx x_0$), we find that the surface potential barrier is not depending directly on the N_s but rather through the value of donor density N_D. The corresponding relationship can be found through the electroneutrality condition

$$qN_s \cdot 2\pi Rl = qN_D \cdot \pi \left(R^2 - (R - x_0)\right)^2,$$

according to which the electrical charge on the surface of the nanowire equals the charge in the depletion region throughout the nanowire, while l denotes the length of the nanowire. This implies

$$x_0 = R - \sqrt{R^2 - 2\frac{N_s}{N_D}R} \qquad \Rightarrow \qquad N_D \approx 2\frac{N_s}{R}$$

if we regard the depletion region with $R \approx x_0$. Therefore the density of bulk donors cannot be considered constant and equal to the density of ionized donors N_D anymore, but varying with the density of occupied surface states. Hence, [Fort et al. 2010] found the density of free electrons in the bulk material n_b as

$$n_b = N_D - 2\frac{N_s}{R}, \tag{2.13}$$

which is the difference between the density of donors and the electrons occupying the surface.

2.5 The Conductance Formulas

Equations (2.10) and (2.12) show clearly that a change in the density of occupied surface states N_s effects the potential barrier and results consequently in a change in conductance. Therefore, in combination with the conductance formulas (2.9) and the relationships for the density of bulk electrons given in (2.11) and (2.13), the equation

$$G(T, V_s) = \begin{cases} G_0 \left(1 - \frac{N_s}{N_D R}\right) \exp\left(\frac{-q^2 N_D}{4\varepsilon kT}\right) + G_C & \text{for } \lambda_D \approx w \\ G_0 \exp\left(-\frac{q^2 N_s^2}{2\varepsilon N_D kT}\right) + G_C & \text{for } \lambda_D \ll w. \end{cases} \tag{2.14}$$

describes the conductance of a sensor, depending on the grade of carrier depletion. The case $\lambda_D \approx w$ marks the completely depleted nanowire and $\lambda_D \ll w$ applies to a nanostructure of greater diameter/width, so that there exists a inner region unaffected of surface phenomena.

The pre-exponential factor G_0 has the form of

$$G_0 = const \cdot q \; v \; n_b, \tag{2.15}$$

while the velocity term v changes depending on the underlying theory. For a partly depleted nanostructures the n_b can be set to the density of donors N_D, according to (2.11). In the case of complete charge depletion the value of n_b was excluded from G_0 and separately considered, as described by equation (2.13).

The additive term G_C in equation (2.14) is a parameter used to provide for various conduction phenomena, like a drift in the sensor signal, and to give the baseline level. This parameter as addition to the conductance formula was introduced by [Ding et al. 2001].

The diffusion theory states the dependence of G_0 on $V_s^{1/2}$. Given the already exponential dependence of G on V_s in the second term of (2.14), this can be neglected. The same reasoning is applicable to the temperature-dependence of the pre-exponential factor, since $\mu_s \propto T^{-3/2}$ according to [Madou and Morrison 1989]. Some authors like Fort et al. have nevertheless chosen to express G_0 as $G_0' T^{-3/2}$ and regard G_0' as a constant. Otherwise, the whole pre-exponential factor G_0 is often regarded as a constant and can therefore be estimated, along with G_C, in a parameter fitting process.

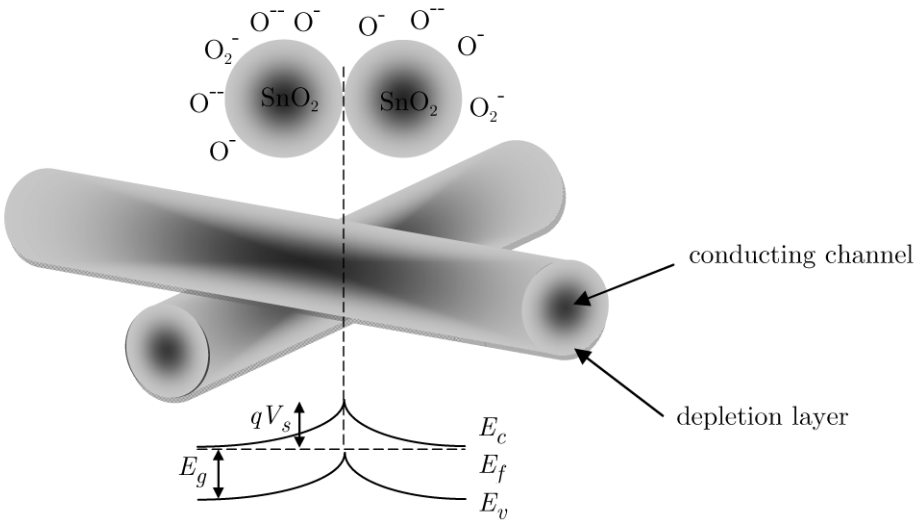


Figure 2.1: Illustration of the conduction in a network of nanowires. The charge depletion region under the surface of the nanowires is due to the adsorption of oxygen species. The resulting potential barrier between touching nanowires is depicted in the band diagram where E_c, E_f, E_v and E_g are the conduction band, Fermi level, valence band and the band gap. qV_s denotes the height of the surface potential barrier. See [Choi et al. 2008], Figure 4 and [Zima 2009], Figure 2.14.

Despite the fact that equations like (2.14) were intended to describe the conductance in layer of semiconducting material, [Fort et al. 2010] argued their applicability to a bundle of semiconducting nanowires. The current path consists of many nanowires with small contact

regions. For this type of sensor the contact points between nanowires act similarly to the contact regions between the grains in a porous layer and therefore build up a surface potential barrier, which electrons have to overcome in order to contribute to the charge flow and add to the sensor conductance. Indeed, according to [Comini et al. 2009], the conduction mechanism of nanowire bundles is dominated by the intercrystalline boundaries at the connection regions between nanowires, as these contact points provide most of the resistance of the sample.

Regarding the modeling of conductance in a single nanowire sensor the same group of authors have shown that also in this case the conduction model (2.14) is applicable, [Fort et al. 2009]. Despite this, the case of a single nanowire, can also be modeled through classical semiconductor transport equations like the Drift-Diffusion Model or the Poisson-Boltzmann equation, [Katterbauer 2010].

Chapter 3

Surface Reaction Models

This section will give surface reaction models derived from chemical reaction equations, which depict the interaction of different kinds of gases and a metal oxide surface, or, to be more specific, on tin oxide (SnO_2).

3.1 Introduction

The ability of gas sensors to detect the presence of chemicals in the atmosphere depends on the interaction between gas and sensor surface. A strong interaction is possible due to the fact that at the surface of the sensing area the periodicity of the crystal lattice is disrupted, thus increasing the reactivity of the sensing area. Also factors like doping, alloying, adsorbates or impurities would influence the reactivity of a surface.

In these cases localized energy states arise at the surface, which are able to exchange electrons with the sensor bulk atoms. These energy levels are called surface states. The fundamental theory to the idea of surface states can be found in the work of [Morrison 1990; Madou and Morrison 1989]. The group Ding et al., who based their work on Morrison et al., differentiate between intrinsic and extrinsic surface states and characterize them as follows:

Intrinsic surface states: They are created by the semiconductor itself, and include energy levels stemming from impurities, doping and oxygen vacancies in the metal oxide lattice. These energy states are the cause for the sensor response to inert gases, like Argon (Ar) in [Ding et al. 2001] and Nitrogen (N_2) in [Lu et al. 2006], where there is no reaction possible between gas and surface.

Extrinsic surface states: These localized surface energy levels, on the other hand, are created through adsorbed gas molecules at the SnO_2 surface, like adsorbed oxygen.

In both cases the exchange of electrons between the conduction band of the metal oxide and a surface state leads to an occupied and consequently charged surface state.

In the following, N_s represents all the occupied surface states, which includes those intrinsically present in the semiconductor, N_{si}, and the extrinsic surface states N_{se}, which essentially consist of all adsorbed gas species on the sensor surface which got ionized by gaining or loosing electrons:

$$N_s = N_{si} + N_{se}.$$

Intrinsic surface states are generally negatively charged because of the reception of a conduction electron. For extrinsic surface states the situation depends entirely on the type of reaction taking place between SnO_2 surface and adsorbed gas species. In the case of an adsorbed oxygen species occupying a surface site the oxygen gets ionized trough an electron stemming from the conduction band. This leads to a negatively charged occupied surface state.

Some reducing gases like CO are also possible to react with the SnO_2 surface directly instead of through an pre-adsorbed oxygen species. This leads to CO acting as a donor for reintroducing an electron into the conduction band of the tin oxide and becoming a positively charged occupied extrinsic surface state.

3.2 Intrinsic Surface States

According to the intrinsic surface state trapping model of [Ding et al. 2001] a temperature increase in an argon atmosphere leads to a thermal excitation of valence electrons. This brings them into the conduction band where they act as free charge carriers and thereby increase the electrical conductivity. The electrons in the conduction band are also thermally excited and some may acquire enough additional energy to overcome the potential barrier at the surface and be trapped in unoccupied surface states.

There are more electrons being trapped in an surface state than there are electrons leaving an occupied surface state. Therefore the potential barrier is building up until a new thermodynamic equilibrium is established.

If the sensor undergoes a quick decrease in temperature the above mentioned process is reversed. Conductance electrons will return to the valence band and occupied surface states will release their electrons back into the conduction band. There will be more electrons evacuating an occupied surface state than electrons being trapped into an unoccupied state. Because of this the potential barrier is lowered again to a new equilibrium.

The effect of a change in temperature effects the thermal excitation of an electron into a different energy band much more quickly than the process of trapping electrons into or releasing them from surface states.

Although the surface states indicate bands of energy levels it is more convenient to handle

them as a single energy level E_t. To evaluate the rates for electron transfer between surface state E_t and conductance band E_c we assume that the rate of electron transfer is of first order. This means that

a) the rate of electron trapping is proportional to the density of unoccupied surface states and the density of electrons in the conduction band at the surface and

b) the rate of electron release to the conduction band is proportional to the density of occupied surface states.

As the sensor surface has no contact to any type of reactive gas the conductivity is uniquely determined by the amount of occupied (i.e., ionized) intrinsic surface states, N_{si}, which are part of all intrinsic surface states N_i — occupied and unoccupied.

The above explained reactions are therefore shown in the reaction equation

$$N_i + e^- \underset{k_{-\mathrm{i}}}{\overset{k_{\mathrm{i}}}{\rightleftharpoons}} N_{si}, \tag{3.1}$$

where the constants k_{i} and $k_{-\mathrm{i}}$ are the reaction rates for electron trapping and releasing, while

N_i ..... total density of occupied, intrinsic surface states,
N_{si} .... density of occupied, intrinsic and ionized surface states.

The application of the law of mass action, [Lund 1965], gives the overall rate for the change of the electron density in the surface states, which is expressed by

$$\frac{dN_{si}}{dt} = k_{\mathrm{i}} n_s \underbrace{(N_i - N_{si})}_{\text{density of unionized intrinsic surface states}} - k_{-\mathrm{i}} N_{si}. \tag{3.2}$$

This denotes the density of occupied intrinsic surface states for the density of electrons n_s in the conduction band at the surface. The above model can also be found in the work of [Fort et al. 2007].

3.3 Extrinsic Surface States

3.3.1 Adsorption of Oxygen (O_2)

Although oxygen is a oxidizing gas and could therefore be incorporated in the below section dedicated to this type of gases, it was decided to provide a full section for the discussion of O_2-driven surface reactions. The reason for this, is that oxygen, as the dominant non-inert component of air, is the reaction determining gas for measurements intended to model the influence of gases on tested gas sensors in an real-life application.

At temperatures between 100 and 600°C oxygen molecules in the atmosphere interact with the SnO_2 surface. At first O_2 is adsorbed to the surface of the metal oxide through physisorption, without influencing its electric properties. With the following chemisorption and ionization the oxygen gets possibly dissociated and bound to the surface through an unoccupied chemisorption site for oxygen in molecular (O_2^-) and atomic (O^-, O^{--}) form, while extracting electrons from the semiconductor to ionize the chemisorbed oxygen. These electrons are free conduction electrons, stemming from ionized donors, which get trapped in a surface state (i.e., the chemisorbed oxygen species) and thereby cannot anymore contribute to the conduction of the sensor.

As the temperature increases the oxygen is adsorbed in the following forms and ways

$$O_{2(\mathrm{gas})} \rightleftharpoons O_{2(\mathrm{ads})} \underset{-e^-}{\overset{+e^-}{\rightleftharpoons}} O^-_{2(\mathrm{ads})} \underset{-e^-}{\overset{+e^-}{\rightleftharpoons}} O^{--}_{2(\mathrm{ads})} \rightleftharpoons 2O^-_{(\mathrm{ads})} \underset{-e^-}{\overset{+e^-}{\rightleftharpoons}} 2O^{--}_{(\mathrm{ads})} \rightleftharpoons 2O^{--}_{(\mathrm{lat})},$$

while the subscripts (*gas*), (*ads*) and (*lat*) denote that the corresponding gas species resides unbounded in the atmosphere, is adsorbed and is bounded into the crystal lattice of the surface respectively.

The chemisorption of atmospheric oxygen starts at temperatures around 100°C. Below that temperature only physisorption, if anything, takes place, leading to $O_{2(\mathrm{ads})}$ on the surface. In the temperature range from 100 to 150°C the dominating oxygen species on the tin oxide surface is chemisorbed molecular oxygen $O^-_{2(\mathrm{ads})}$, which later may acquire another electron to form $O^{--}_{(\mathrm{ads})}$ and dissociate in higher temperatures. Above 150°C the atmospheric oxygen dissociates upon contact and adsorbed in atomic form ($O^-_{(\mathrm{ads})}$ and $O^{--}_{(\mathrm{ads})}$), while the singly ionized form dominates the temperature range below 450°C. Above this temperature O^{--} is predominant. This species is then directly incorporated as bridging oxygen into the SnO_2 lattice in the form of $O^{--}_{(\mathrm{lat})}$ above 600°C, according to [Tabata et al. 2003].

Corresponding to [Bârsan and Weimar 2001] and [Fort et al. 2006b], the following two-step *reaction equation* for oxygen illustrates the reversible surface reaction

$$\text{Chemisorption:} \qquad \frac{\beta}{2} O_{2(\mathrm{gas})} + S \underset{k_{-\mathrm{O}}}{\overset{k_{\mathrm{O}}}{\rightleftharpoons}} O_{\beta(\mathrm{ads})}, \tag{3.3a}$$

$$\text{Ionization:} \qquad O_{\beta(\mathrm{ads})} + \alpha \cdot e^- \underset{k_{-\mathrm{O}^-}}{\overset{k_{\mathrm{O}^-}}{\rightleftharpoons}} O^{-\alpha}_{\beta(\mathrm{ads})}, \tag{3.3b}$$

where k_{O_β} and $k_{-\mathrm{O}_\beta}$ are the rate constants for the chemisorption and $k_{\mathrm{O}_\beta^{-\alpha}}$ and $k_{-\mathrm{O}_\beta^{-\alpha}}$ the rate constants corresponding to the oxygen ionization. The used notations are:

- $O_{2(\mathrm{gas})}$ ..... oxygen molecule in the ambient atmosphere and their concentration $[O_{2(\mathrm{gas})}]$,
- e^- ..... electrons, which can reach the surface (with concentration $n_s = [e^-]$),
- S ..... unoccupied chemisorption site,

$O_{\beta(\text{ads})}$ ..... chemisorbed oxygen species occupying a chemisorption site for oxygen at the surface (their concentration is denoted as $N_{O_\beta} = [O_{\beta(\text{ads})}]$),
$O^{-\alpha}_{\beta(\text{ads})}$ ..... chemisorbed and ionized oxygen species with concentration $N_{O_\beta^{-\alpha}}$,
S_t ..... unoccupied or occupied chemisorption site on the surface,

while

$$\alpha = \begin{cases} 1 & \text{for singly ionized forms,} \\ 2 & \text{for doubly ionized forms} \end{cases} \quad \text{and} \quad \beta = \begin{cases} 1 & \text{for atomic forms,} \\ 2 & \text{for molecular forms.} \end{cases}$$

As above mentioned, for temperature ranges below 150°C the molecular oxygen species are dominant ($\beta = 2$), while above this temperature oxygen chemisorbs in atomic form, both in singly ionized form ($\alpha = 1$). At elevated temperatures around 400°C the doubly ionized oxygen species ($\alpha = 2$ and $\beta = 1$) is predominant.

The presence of these species leads to the formation of a depletion layer and subsequent a space-charge region at the surface of the tin oxide. This region then leads to a potential barrier at the surface of the semiconductor. This process decreases the conductance of SnO_2, depending on the density of the chemisorbed surface oxygen on the semiconductor surface, which itself depends on the partial pressure or the concentration of oxygen in the atmosphere.

Rate Equation

The activation energies for adsorption and desorption included in the reaction constants, k_{O_β}, k_{-O_β}, $k_{O_\beta^{-\alpha}}$ and $k_{-O_\beta^{-\alpha}}$ as well as the mass action law applied to (3.3) yield

$$k_{O_\beta}([S_t] - N_{O_\beta} - N_{O_\beta^{-\alpha}})[O_{2(\text{gas})}]^{\beta/2} = k_{-O_\beta} N_{O_\beta}, \tag{3.4a}$$

$$k_{O_\beta^{-\alpha}} n_s^\alpha N_{O_\beta} = k_{-O_\beta^{-\alpha}} N_{O_\beta^{-\alpha}}, \tag{3.4b}$$

while a first order adsorption reaction is assumed, which complies to the majority of literature regarding the derivation of surface reaction models, [Fort et al. 2006b; Ding et al. 2001].

On these bases we get the *rate equations* in the form of differential equations, describing the rate of adsorbed (neutral and ionized) oxygen density change, which read like

$$\frac{dN_{O_\beta}}{dt} = k_{O_\beta}([S_t] - N_{O_\beta} - N_{O_\beta^{-\alpha}})[O_{2(\text{gas})}]^{\beta/2} - k_{-O_\beta} N_{O_\beta} - \frac{dN_{O_\beta^{-\alpha}}}{dt}, \tag{3.5a}$$

$$\frac{dN_{O_\beta^{-\alpha}}}{dt} = k_{O_\beta^{-\alpha}} n_s^\alpha N_{O_\beta} - k_{-O_\beta^{-\alpha}} N_{O_\beta^{-\alpha}}. \tag{3.5b}$$

3.3.2 The influence of Humidity (H_2O)

Humidity is an ubiquitous factor for the operation of gas sensors in ambient air, therefore the influence of gaseous H_2O on the sensor response should be analyzed. At temperatures between 100 and 500°C, the interaction of a metal oxide surface with water vapor leads to molecular water and hydroxyl groups adsorption, although above 200°C no more molecular water can be found on a SnO_2 surface.

There are three types of mechanisms explaining the experimentally proven (as seen in [Bârsan and Weimar 2001]) increase of surface conductivity in the presence of water vapor. All of them take into account that

a) water vapor increases surface conductance and

b) the effect is reversible.

Nevertheless, after studying the adsorption mechanism of CO under the influence of humidity, [Bârsan and Weimar 2003] strongly suggests that the reaction mechanism

$$H_2O_{(gas)} + Sn_{(lat)} + O_{(lat)} \underset{k_{-H_2O}}{\overset{k_{H_2O}}{\rightleftharpoons}} (Sn^+_{(lat)} - OH^-_{(ads)}) + (OH)^+_{(lat)} + e^-$$

is chosen, where $(Sn^+ - OH^-)$ is a so called isolated OH group and $OH^+_{(lat)}$ a rooted one. The electron on the right hand side, which is subsequently injected into the conduction band, stems from the rooted hydroxyl group as it gets ionized and becomes a donor.

To derive a rate equation based surface reaction model we simplify the above reaction equation such that the lattice oxygen and tin species are treated as unoccupied surface sites $[S]$. This consideration leads to

$$H_2O_{(gas)} + 2S \underset{k_{-H_2O}}{\overset{k_{H_2O}}{\rightleftharpoons}} OH^-_{(ads)} + H^+_{(ads)} + e^-. \tag{3.6}$$

If we would anticipate chemical reactions between the hydroxyl groups and other adsorbing gas species we give separate rate equations for $[OH^-_{(ads)}]$ and $[H^+_{(ads)}]$, but otherwise those two adsorbed species can be represented through the joint rate equation

$$\frac{d[OH^-_{(ads)} + H^+_{(ads)}]}{dt} = k_{H_2O}[H_2O][S]^2 - k_{-H_2O} n_s [OH^-_{(ads)} + H^+_{(ads)}]. \tag{3.7}$$

The link to other reaction models would be provided by the equation for unoccupied surface sites $[S] = [S_t] - [OH^-_{(ads)} + H^+_{(ads)}]$ minus the concentration of various other adsorbed gas species.

To avoid such a complicated expansion of existing surface reaction models, the effect of humidity on the operation of a metal oxide gas sensor can be circumvented by only taking measurements in a dry atmosphere. Should one nevertheless want to quantify the effect of water adsorption on the charge carrier concentration, n_s (which is normally proportional to the measured conductance), one could include the effect of water by considering the effect of an increased background of free charge carriers on the adsorption of varying gases. This of course is a simplification that would work if the respective test gas is not probable to react with OH^- groups on the surface, as, for example, CO clearly is [Zima 2009]. To find humidity as a reaction product from surface reactions of hydrogen-containing test gases is also possible. As a consequence all subsequent surface reaction models are intended for a dry atmosphere.

We will now turn our attention to surface reactions stemming from various gas species. The mechanism of gas detection is usually intimately related to reactions between a target gas and previously adsorbed oxygen species, although a direct adsorption of gas species onto the surface is also possible for various gas species. Independent of the use of pre-adsorbed gas species as an intermediary step, gases can be classified into two major groups, depending on their mode of operation in a reduction-oxidation reaction. The chemical reactions on a SnO_2 surface can therefore by divided into two basic cases, depending on the type of target gas causing it:

Reducing Gases: Gases acting as reducing agents in a redox reaction generally are electron donors. These gas species therefore increase the conductance of the semiconductor upon adsorption, by releasing electrons into the conduction band of the sensor.

Oxidizing Gases: Oxidizing gases act as an oxidizing agent in a redox reaction and operate in a reversed manner to a reducing gas by becoming an electron acceptor. This type of gases cause a decrease in the semiconductor conductance by binding free electrons into unoccupied surface states.

3.3.3 Reducing Gases

It is considered in [Fort et al. 2006b] that, if the reducing gas concentration $[Red_{(gas)}]$ is low with respect to the oxygen concentration in the carrier gas (at most 400 ppm (i.e., 0,04%) $[Red_{(gas)}]$ versus 2.1×10^5 ppm (i.e., 21%) $[O_2]$), then the most probable reaction would be: A reducing gas $Red_{(gas)}$ reacts with the chemisorbed oxygen $O^-_{(ads)}$ in the atomic and singly charged form ($\alpha = \beta = 1$) on the semiconductor surface, releasing electrons into the semiconductor and desorbing as the gaseous reaction product $RedO_{(gas)}$ from the semiconductor surface, [Häusler

2004]:

$$\mathrm{Red}_{(\mathrm{gas})} + \mathrm{O}^-_{(\mathrm{ads})} \xrightarrow{k_{\mathrm{RedO}}} \mathrm{RedO}_{(\mathrm{gas})} + S + e^- \tag{3.8}$$

with S a vacant chemisorption site, formerly occupied by the oxygen reacting with the reduced gas and k_{RedO} the rate constant for the oxidation reaction (i.e., the rate for the $\mathrm{RedO}_{(\mathrm{gas})}$ production). In this case the reaction is irreversible, as shown by the reaction arrow.

The release of the charge carriers into the conduction band of the semiconductor leads to an increase in conductance on the semiconductor surface, depending on the concentration of reducing gas in the atmosphere.

We modify the rate equations of oxygen adsorption (3.5) to consider the reaction between the reducing gas and the ionized oxygen, which leads us to

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}([S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^-})p_{\mathrm{O}_2}^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.9a}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-}n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-}N_{\mathrm{O}^-} - \frac{[\mathrm{RedO}_{(\mathrm{gas})}]}{dt}, \tag{3.9b}$$

$$\frac{[\mathrm{RedO}_{(\mathrm{gas})}]}{dt} = k_{\mathrm{RedO}}[\mathrm{Red}_{(\mathrm{gas})}]N_{\mathrm{O}^-}. \tag{3.9c}$$

As this model describes the most basic of possible reactions caused by a reducing gas, it can be implemented to model the sensor response to all reducing gases, or more generally, gases ultimately increasing the sensor conductance. But, as even surface reactions caused by a well researched gas like CO are quiet varied, a sensor model derived from the actual chemical reactions caused by a specific gas can be much more detailed.

Carbon Monoxide (CO)

Carbon Monoxide is a highly inflammable gas which occurs in combustion motors as well as a side product through the burning of fossil fuels. A careful screening of CO levels is required not only because carbon monoxide is a respiratory poison in higher quantities, but also to monitor the combustion efficiency and pollutant emission in various industrial settings.

CO is considered to react with pre-adsorbed oxygen or lattice oxygen if there are no oxygen adsorbates. If there is no pre-adsorbed oxygen on the surface of the metal oxide, the CO reacts with lattice oxygen and donates electrons, therefore increasing the surface conductivity, according to the reaction equation

$$\text{Chemisorption:} \qquad \mathrm{CO}_{(\mathrm{gas})} + S \underset{k_{-\mathrm{CO}}}{\overset{k_{\mathrm{CO}}}{\rightleftharpoons}} \mathrm{CO}_{(\mathrm{ads})}, \tag{3.10a}$$

Ionization: $$\mathrm{CO}_{(\mathrm{ads})} \underset{k_{-\mathrm{CO}^+}}{\overset{k_{\mathrm{CO}^+}}{\rightleftharpoons}} \mathrm{CO}^+_{(\mathrm{ads})} + e^-. \tag{3.10b}$$

This yields the *rate equations* for CO, which describe the sensor dynamics in the presence of carbon monoxide:

$$\frac{dN_{\mathrm{CO}}}{dt} = k_{\mathrm{CO}}[S][\mathrm{CO}] - k_{-\mathrm{CO}}N_{\mathrm{CO}} - \frac{dN_{\mathrm{CO}^+}}{dt}, \tag{3.11a}$$

$$\frac{dN_{\mathrm{CO}^+}}{dt} = k_{\mathrm{CO}^+}N_{\mathrm{CO}} - k_{-\mathrm{CO}^+}n_s N_{\mathrm{CO}^+}, \tag{3.11b}$$

for $[S] = [S_t] - N_{\mathrm{CO}} - N_{\mathrm{CO}^+}$.

Equations (3.10) refer to the case of CO measurements in an atmosphere consisting of inert gases. However in an air atmosphere, which indicates the presence of adsorbed oxygen at the surface, CO increases the surface conductance, according to [Bârsan and Weimar 2001]. This phenomenon is explained in

$$\mathrm{CO}_{(\mathrm{gas})} + \mathrm{O}^-_{(\mathrm{ads})} \xrightarrow{k_{\mathrm{CO}_2}} \mathrm{CO}_{2(\mathrm{gas})} + e^- + S, \tag{3.12}$$

with the reaction constant k_{CO_2} for carbon dioxide production.

The application of the law of mass action gives us

$$\frac{d[\mathrm{CO}_{2(\mathrm{gas})}]}{dt} = k_{\mathrm{CO}_2}p_{\mathrm{CO}}N_{\mathrm{O}^-},$$

the corresponding rate equation for CO, which can be added to equations (3.3) in the manner described in (3.9).

According to [Hahn et al. 2003], depending on the amount of oxygen in the atmosphere, both of the reaction mechanisms described above can occur, especially in low concentrations of O_2 (about 25 to 50 ppm or 0.0025 to 0.005%) as opposed to 250ppm CO (i.e., 0.025 %). However the higher the amount of oxygen in the atmosphere and therefore in adsorbed and ionized form on the SnO_2 surface, the more likely is reaction (3.12). A combination of these two reaction mechanisms for CO and the rate equations for oxygen leads to

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S][\mathrm{O}_2]^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt} \tag{3.13a}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-}n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-}N_{\mathrm{O}^-} - \frac{d[\mathrm{CO}_{2(\mathrm{gas})}]}{dt} \tag{3.13b}$$

$$\frac{d[\mathrm{CO}_{2(gas)}]}{dt} = k_{\mathrm{CO}_2}[\mathrm{CO}]N_{\mathrm{O}^-} \tag{3.13c}$$

$$\frac{dN_{\mathrm{CO}}}{dt} = k_{\mathrm{CO}}[S][\mathrm{CO}] - k_{-\mathrm{CO}}N_{\mathrm{O}} - \frac{dN_{\mathrm{CO}^+}}{dt} \tag{3.13d}$$

$$\frac{dN_{\mathrm{CO}^+}}{dt} = k_{\mathrm{CO}^+}N_{\mathrm{CO}} - k_{-\mathrm{CO}^+}n_s N_{\mathrm{CO}^+} \tag{3.13e}$$

with $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^-} - N_{\mathrm{CO}} - N_{\mathrm{CO}^+})$, the concentration of unoccupied surface states.

Molecular Hydrogen (H_2)

Hydrogen is a highly combustible gas and is regarded as a promising candidate as future energy carrier for rocket fuel or fuel cells. It is also a contaminating component in chemical industries, which necessitates the reliable detection of hydrogen.

At comparatively low temperatures (100 to 300°C) there is no dissociation of H_2, according to [Gong et al. 2004]. Therefore, the hydrogen molecules react directly with the adsorbed oxygen species on the surface. In this range of temperature the adsorbed oxygen is comprised of molecular and atomic singly ionized species (O_2^- and O^-). As already mentioned, for temperatures below 150°C the molecular oxygen species (O_2^-) is dominant on the sensor surface. Therefore reaction

$$2\mathrm{H}_{2(\mathrm{gas})} + \mathrm{O}_2^- \xrightarrow{k_{\mathrm{H}_2(1)}} 2\mathrm{H}_2\mathrm{O}_{(\mathrm{gas})} + S + e^- \tag{3.14}$$

is likely to occur. In combination with the reaction kinetics for oxygen adsorption it leads to

$$\frac{dN_{\mathrm{O}_2}}{dt} = k_{\mathrm{O}_2}[S][\mathrm{O}_2] - k_{-\mathrm{O}_2}N_{\mathrm{O}_2} - \frac{dN_{\mathrm{O}_2^-}}{dt}, \tag{3.15a}$$

$$\frac{dN_{\mathrm{O}_2^-}}{dt} = k_{\mathrm{O}_2^-}n_s N_{\mathrm{O}_2} - k_{-\mathrm{O}_2^-}N_{\mathrm{O}_2^-} - \frac{d[\mathrm{H}_2\mathrm{O}_{(\mathrm{gas})}]}{dt}, \tag{3.15b}$$

$$\frac{d[\mathrm{H}_2\mathrm{O}_{(\mathrm{gas})}]}{dt} = k_{\mathrm{H}_2(1)}[\mathrm{H}_2]N_{\mathrm{O}_2^-}, \tag{3.15c}$$

while $[S] = ([S_t] - N_{\mathrm{O}_2} - N_{\mathrm{O}_2^-})$.

In a temperature regime above 150°C the singly charged atomic oxygen species (O^-) is dominant and makes reaction

$$\mathrm{H}_{2(gas)} + \mathrm{O}^- \xrightarrow{k_{\mathrm{H}_2(2)}} \mathrm{H}_2\mathrm{O}_{(gas)} + S + e^- \tag{3.16}$$

more likely to occur. If we derive the corresponding rate equation and combine it with the reaction kinetics of oxygen adsorption (in O^- form) we get

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S][\mathrm{O}_2]^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.17a}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-}n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-}N_{\mathrm{O}^-} - \frac{d[\mathrm{H}_2\mathrm{O}_{(\mathrm{gas})}]}{dt}, \tag{3.17b}$$

$$\frac{d[\mathrm{H}_2\mathrm{O}_{(\mathrm{gas})}]}{dt} = k_{\mathrm{H}_2(2)}[\mathrm{H}_2]N_{\mathrm{O}^-} \tag{3.17c}$$

with $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^-})$ the concentration of unoccupied surface states. This second

reaction mechanism for H_2 was also employed to derive a response model in the work of [Yamazoe and Shimanoe 2010].

The work of [Malyshev and Pislyakov 2008], on the other hand, suggests that the adsorption of molecular hydrogen onto a metal oxide surface in temperatures above 350 to 400°C follows a two step reaction. While the first part of this mechanism consists of the dissociation of H_2 and the production of OH groups, the second part consists of the recombination of these species leading to the desorption of gaseous H_2O. As the temperature is above 350°C the adsorbed oxygen species on the surface exist in the form of O^{--}. The associated reaction mechanism leads to

$$\begin{aligned}
&\frac{1}{2}H_{2(\text{gas})} \rightleftharpoons H_{(\text{gas})},\\
&H_{(\text{gas})} + O^{--}_{(\text{ads})} \rightarrow OH^{-}_{(\text{ads})} + e^{-},\\
&H_{(\text{gas})} + OH^{-}_{(\text{ads})} \rightarrow H_2O_{(\text{gas})} + S + e^{-},
\end{aligned}$$

which can be combined into one reaction, leading to the overall reaction equation

$$H_{2(\text{gas})} + O^{--}_{(\text{ads})} \xrightarrow{k_{H_2(3)}} H_2O_{(\text{gas})} + S + 2e^{-}. \tag{3.19}$$

This combined reaction mechanism coincides with the reaction equation pertaining to H_2 and O^{--} given by [Gong et al. 2004]. The corresponding rate equation can be combined with the reaction kinetics for O^{--}, leading to

$$\frac{dN_O}{dt} = k_O[S][O_2]^{1/2} - k_{-O}N_O - \frac{dN_{O^{--}}}{dt}, \tag{3.20a}$$

$$\frac{dN_{O^{--}}}{dt} = k_{O^{--}}n_s^2 N_O - k_{-O^{--}}N_{O^{--}} - \frac{d[H_2O_{(\text{gas})}]}{dt}, \tag{3.20b}$$

$$\frac{d[H_2O_{(\text{gas})}]}{dt} = k_{H_2(3)}[H_2]N_{O^{--}}, \tag{3.20c}$$

while $[S]$ denotes the available surface sites and are given by $[S_t] - N_O - N_{O^{--}}$.

Methane (CH_4)

Methane is formed through the bacterial decomposition of organic material and the production of fossil fuels and acts as a greenhouse gas. The accurate detection of CH_4 is therefore especially important for environmental control and industrial production processes.

While the possible chemical reactions taking place between Methane and a metal oxide surface are numerous, are diverse in their reaction products (hydrogen, gaseous water and various carbon-hydrogen-oxygen compounds [Kohl 2001]) and are not completely understood, the reaction mechanisms proposed by [Quaranta et al. 1999] were chosen to form the surface

reaction model

$$\mathrm{CH_{4(gas)}} + 2\mathrm{O_{(lat)}} \xrightarrow{k_{\mathrm{CH_4}(1)}} \mathrm{CH_3^-O_{(lat)}} + \mathrm{H^-O_{(lat)}}, \tag{3.21a}$$

$$\mathrm{CH_3^-O_{(lat)}} + 2\mathrm{O_{(lat)}} + 2\mathrm{O^-_{(ads)}} \xrightarrow{k_{\mathrm{CH_4}(2)}}, \mathrm{CO_{2(gas)}} + 3\mathrm{H^-O_{(lat)}} + 2e^- \tag{3.21b}$$

$$2\mathrm{CH_3^-O_{(lat)}} + \mathrm{O^-_{(ads)}} \xrightarrow{k_{\mathrm{CH_4}(3)}} .\mathrm{H_2O_{(gas)}} + \mathrm{C_2H_{4(gas)}} + 2\mathrm{O_{(lat)}} + e^- \tag{3.21c}$$

The lattice oxygen occurring in these reactions is treated as an unoccupied adsorption site $[S]$, as $\mathrm{O_{(lat)}}$ is not consumed by the desorption of gases to form a oxygen vacancy. Even more complex reactions can be found in [Kohl 2001].

After the application of the law of mass action the rate equations

$$\frac{d[\mathrm{CH^-_{3(ads)}}]}{dt} = k_{\mathrm{CH_4}(1)}[S][\mathrm{CH_4}] - \frac{d[\mathrm{CO_{2(gas)}}]}{dt} - \frac{d[\mathrm{H_2O_{(gas)}}]}{dt}, \tag{3.22a}$$

$$\frac{d[\mathrm{CO_{2(gas)}}]}{dt} = k_{\mathrm{CH_4}(2)}[\mathrm{CH^-_{3(ads)}}]N^2_{\mathrm{O^-}}, \tag{3.22b}$$

$$\frac{d[\mathrm{H_2O_{(gas)}}]}{dt} = k_{\mathrm{CH_4}(3)}[\mathrm{CH^-_{3(ads)}}]N_{\mathrm{O^-}} \tag{3.22c}$$

can be derived from (3.21).

As the contributing oxygen species are of the molecular, singly ionized form, the temperature range in the case of reactions (3.21), would be above 150°C. Therefore in combination with the rate equations for oxygen adsorption we get

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S]p^{1/2}_{\mathrm{O_2}} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O^-}}}{dt}, \tag{3.23a}$$

$$\frac{dN_{\mathrm{O^-}}}{dt} = k_{\mathrm{O^-}}n_s N_{\mathrm{O}} - k_{-\mathrm{O^-}}N_{\mathrm{O^-}} - \frac{d[\mathrm{CO_{2(gas)}}]}{dt} - \frac{d[\mathrm{H_2O_{(gas)}}]}{dt}, \tag{3.23b}$$

$$\frac{d[\mathrm{CH^-_{3(ads)}}]}{dt} = k_{\mathrm{CH_4}(1)}[S]p_{\mathrm{CH_4}} - \frac{d[\mathrm{CO_{2(gas)}}]}{dt} - \frac{d[\mathrm{H_2O_{(gas)}}]}{dt}, \tag{3.23c}$$

$$\frac{d[\mathrm{CO_{2(gas)}}]}{dt} = k_{\mathrm{CH_4}(2)}[\mathrm{CH^-_{3(ads)}}]N^2_{\mathrm{O^-}}, \tag{3.23d}$$

$$\frac{d[\mathrm{H_2O_{(gas)}}]}{dt} = k_{\mathrm{CH_4}(3)}[\mathrm{CH^-_{3(ads)}}]N_{\mathrm{O^-}}, \tag{3.23e}$$

as reaction model for the methane adsorption, while $[S] = [S_t] - N_{\mathrm{O}} - N_{\mathrm{O^-}} - [\mathrm{CH^-_{3(ads)}}]$. In these reactions hydrogen was involved as an reaction product, therefore subsequent reactions fueled by adsorbed hydrogen may take place on the sensor surface. The production of humidity could also influence the operation of the sensor in different ways than the reaction mechanisms described above. It is debatable if these two factors should actually be considered when devising a sensor model, as the resulting equations would be much more extensive.

Hydrogen Sulfide (H_2S)

Hydrogen sulfide, or reaction products stemming from its combination with other compounds, is highly toxic to humans and also acts as an environmental pollutant. Therefore the detection of this flammable gas is very important for environmental control.

H_2S is known to react with pre-adsorbed oxygen in the manner of a typical reducing gas, once it comes in contact with a SnO_2 surface. [Liu et al. 2009] proposed different reaction mechanisms, depending on the oxygen species currently predominating the sensor surface.

In temperatures below 150°C the dominating molecular oxygen species O_2^- leads to

$$H_2S_{(gas)} + O^-_{2(ads)} \xrightarrow{k_{H_2S(1)}} SO_{2(gas)} + H_{2(gas)} + e^- + S \tag{3.24}$$

as a probable reaction mechanism. The corresponding rate equations, combined with the reaction kinetics of oxygen adsorption of the molecular species, are

$$\frac{dN_{O_2}}{dt} = k_{O_2}[S][O_2] - k_{-O_2}N_{O_2} - \frac{dN_{O_2^-}}{dt}, \tag{3.25a}$$

$$\frac{dN_{O_2^-}}{dt} = k_{O_2^-}n_sN_{O_2} - k_{-O_2^-}N_{O_2^-} - \frac{d[SO_{2(gas)}]}{dt}, \tag{3.25b}$$

$$\frac{d[SO_{2(gas)}]}{dt} = k_{H_2S(1)}[H_2S]N_{O_2^-}, \tag{3.25c}$$

while the concentration of unoccupied surface sites $[S]$ equates $[S_t] - N_{O_2} - N_{O_2^-}$.

The O^- species on the surface at temperatures above 150°C react with H_2S according to

$$H_2S_{(gas)} + 2O^-_{(ads)} \xrightarrow{k_{H_2S(2)}} SO_{2(gas)} + H_{2(gas)} + 2e^- + 2S. \tag{3.26}$$

When included into the reaction kinetics of oxygen adsorption it reads

$$\frac{dN_O}{dt} = k_O[S][O_2]^{1/2} - k_{-O}N_O - \frac{dN_{O^-}}{dt}, \tag{3.27a}$$

$$\frac{dN_{O^-}}{dt} = k_{O^-}n_sN_O - k_{-O^-}N_{O^-} - \frac{d[SO_{2(gas)}]}{dt}, \tag{3.27b}$$

$$\frac{d[SO_{2(gas)}]}{dt} = k_{H_2S(2)}[H_2S](N_{O^-})^2, \tag{3.27c}$$

while $[S] = ([S_t] - N_O - N_{O^-})$.

In an elevated temperature range, where the doubly ionized oxygen species O^{--} is more strongly present at the surface, the equation describing the interaction with H_2S can be written as

$$H_2S_{(gas)} + 2O^{--}_{(ads)} \xrightarrow{k_{H_2S(3)}} SO_{2(gas)} + H_{2(gas)} + 4e^- + 2S, \tag{3.28}$$

which leads to the reaction rate equations

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S][\mathrm{O}_2]^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^{--}}}{dt}, \tag{3.29a}$$

$$\frac{dN_{\mathrm{O}^{--}}}{dt} = k_{\mathrm{O}^{--}}n_s^2 N_{\mathrm{O}} - k_{-\mathrm{O}^{--}}N_{\mathrm{O}^{--}} - \frac{d[\mathrm{SO}_{2(\mathrm{gas})}]}{dt}, \tag{3.29b}$$

$$\frac{d[\mathrm{SO}_{2(\mathrm{gas})}]}{dt} = k_{\mathrm{H_2S}(3)}[\mathrm{H_2S}](N_{\mathrm{O}^{--}})^2, \tag{3.29c}$$

while $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^{--}})$.

In the work of [Malyshev and Pislyakov 1998] and [Kersen and Holappa 2006] the authors proposed another chemical reaction equation as the driving force behind the interaction of H_2S and a metal oxide surface. The proposed reaction mechanism

$$\mathrm{H_2S_{(gas)}} + 3\mathrm{O^{--}_{(ads)}} \xrightarrow{k_{\mathrm{H_2S}(4)}} \mathrm{SO_{2(gas)}} + \mathrm{H_2O_{(gas)}} + 6e^- + 2S \tag{3.30}$$

includes the doubly ionized pre-adsorbed oxygen species O^{--} and takes place in a temperature range well above 150°C. The application of the Law of Mass action and the formulation of the corresponding rate equation leads to

$$\frac{d[\mathrm{SO}_{2(\mathrm{gas})}]}{dt} = k_{\mathrm{H_2S}}[\mathrm{H_2S}](N_{\mathrm{O}^{--}})^3.$$

In combination with the rate equations for the adsorption of O^{--} this leads to

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S][\mathrm{O}_2]^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^{--}}}{dt}, \tag{3.31a}$$

$$\frac{dN_{\mathrm{O}^{--}}}{dt} = k_{\mathrm{O}^{--}}n_s^2 N_{\mathrm{O}} - k_{-\mathrm{O}^{--}}N_{\mathrm{O}^{--}} - \frac{d[\mathrm{SO}_{2(\mathrm{gas})}]}{dt}, \tag{3.31b}$$

$$\frac{d[\mathrm{SO}_{2(\mathrm{gas})}]}{dt} = k_{\mathrm{H_2S}(4)}[\mathrm{H_2S}](N_{\mathrm{O}^{--}})^3, \tag{3.31c}$$

while $[S] = [S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^{--}}$.

3.3.4 Oxidizing Gases

If a oxidizing gas $\mathrm{Oxi_{(gas)}}$ reacts with the surface of a SnO_2 semiconductor the basic reaction is given by

$$\mathrm{Oxi_{(gas)}} + S + e^- \xrightarrow{k_{\mathrm{Oxi}^-}} \mathrm{Oxi^-_{(ads)}}, \tag{3.32}$$

where the oxidizing gas acts as an electron acceptor. Of course there may occur many more complex interactions, depending on the oxidizing gas itself. The corresponding more complicated reaction equations can be found in the subsections dedicated to the individual oxidizing

gases.

In the above basic equation (3.32) k_{Oxi} denotes the rate constant for the $\mathrm{Oxi}^-_{(\mathrm{ads})}$ production. Once again a reversion of this reaction is not likely, as denoted by the one sided arrow. The electron needed for this reaction is taken from the conduction band of the SnO_2 bulk material, thus decreasing the overall conductance.

The application of the law of mass action leads to the rate equation for oxidizing gases, which, in combination with the rate equations corresponding to oxygen (3.5), leads to

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[\mathrm{O}_2]^{1/2}[S] - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.33a}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-}N_{\mathrm{O}} - k_{-\mathrm{O}^-}n_s N_{\mathrm{O}^-}, \tag{3.33b}$$

$$\frac{[\mathrm{Oxi}^-_{(\mathrm{ads})}]}{dt} = k_{\mathrm{Oxi}}[\mathrm{Oxi}]n_s[S]. \tag{3.33c}$$

Here once again the factor $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O}^-} - [\mathrm{Oxi}^-_{(\mathrm{ads})}])$ denotes the number of available unoccupied surface sites for chemisorption, with $[S_t]$ the total number of chemisorption sites and $N_{\mathrm{O}}, N_{\mathrm{O}^-}$ and $[\mathrm{Oxi}^-_{(\mathrm{ads})}]$ the concentration of occupied surface states generated through adsorbed oxygen in neutral and negatively charged form and the adsorbed ionized species $\mathrm{Oxi}^-_{(\mathrm{ads})}$.

Nitrogen Dioxide (NO_2)

As NO_2 gas is one of the most dangerous air pollutants responsible for ozone and acid rain its detection and screening is very important for environmental purposes.

The tin oxide surface reactions in the presence of nitrogen dioxide are somewhat more complex, as there are several possibilities of interaction between the sensor surface and the oxidizing gas not exclusively depending on temperature variation.

At temperatures below 200°C the dominating oxygen species on the SnO_2 surface is O_2^-, which NO_2 is unlikely to react with in a direct way, according to [Ruhland et al. 1998] and [Ionescu et al. 2003]. The NO_2 molecules therefore react directly with the surface tin ions and get ionized themselves:

$$\text{Ionosorption:} \qquad \mathrm{NO}_{2(\mathrm{gas})} + S + e^- \xrightarrow{k_{\mathrm{NO}_2(1)}} \mathrm{NO}^-_{2(\mathrm{ads})}, \tag{3.34a}$$

$$\text{Disintegration:} \qquad \mathrm{NO}^-_{2(\mathrm{ads})} \xrightarrow{k_{\mathrm{NO}_2(2)}} \mathrm{NO}_{(\mathrm{gas})} + \mathrm{O}^-_{(\mathrm{ads})}. \tag{3.34b}$$

The electron in reaction (3.34a) can originate not only from the conduction band of the SnO_2 bulk material but also from an already ionized O_2^- species as the NO_2 molecules forms surface acceptor levels deeper than surface oxygen ions. This reaction leads to an decrease in

conductivity as the height of the surface potential barrier is increased.

A direct reversal of reaction (3.34a) is unlikely but the adsorbed and ionized NO_2 molecules can be disintegrated into a desorbed NO and an ionized surface oxygen ion, as seen in equation (3.34b). Because of this interaction the surface Fermi energy level is increased and the height of the potential barrier at the surface is lowered leading to an increase in conductance.

By applying the law of mass action to the above equations we get

$$\frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} = \underbrace{k_{\mathrm{NO}_2(1)} n_s [\mathrm{NO}_2][S]}_{\text{Ionosorption}} - \underbrace{k_{\mathrm{NO}_2(2)} [\mathrm{NO}^-_{2(\mathrm{ads})}]}_{\text{Disintegration of NO}_2},$$

with the concentration of unoccupied surface states $[S]$.

Combining them with the rate equations for oxygen adsorption (in the molecular oxygen species in this case) leads to the reaction kinetics

$$\frac{dN_{\mathrm{O}_2}}{dt} = k_{\mathrm{O}_2}[\mathrm{O}_2][S] - k_{-\mathrm{O}_2} N_{\mathrm{O}_2} - \frac{dN_{\mathrm{O}_2^-}}{dt}, \tag{3.35a}$$

$$\frac{dN_{\mathrm{O}_2^-}}{dt} = k_{\mathrm{O}_2^-} n_s N_{\mathrm{O}_2} - k_{-\mathrm{O}_2^-} N_{\mathrm{O}_2^-}, \tag{3.35b}$$

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[\mathrm{O}_2]^{1/2}[S] - k_{-\mathrm{O}} N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.35c}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-} n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-} N_{\mathrm{O}^-} + k_{\mathrm{NO}_2(2)} [\mathrm{NO}^-_{2(\mathrm{ads})}], \tag{3.35d}$$

$$\frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} = k_{\mathrm{NO}_2(1)} n_s [\mathrm{NO}_2][S] - k_{\mathrm{NO}_2(2)} [\mathrm{NO}^-_{2(\mathrm{ads})}], \tag{3.35e}$$

where the concentration of unoccupied surface sites $[S]$ is denoted by $([S_t] - N_{\mathrm{O}_2} - N_{\mathrm{O}_2^-} - N_{\mathrm{O}} - N_{\mathrm{O}^-} - [\mathrm{NO}^-_{2(\mathrm{ads})}])$.

It is debatable if the attribution of the NO_2 dissociation to the concentration of O^- should be taken into account, as the molecular oxygen species is dominant in this temperature range.

According to [Leo et al. 1999], the adsorption of NO_2 onto the semiconductor surface can go hand in hand with the disintegration of O_2^-:

$$\mathrm{NO}_{2(\mathrm{gas})} + \mathrm{O}^-_{2(\mathrm{ads})} + 2S + 2e^- \xrightarrow{k_{\mathrm{NO}_2(3)}} \mathrm{NO}^-_{2(\mathrm{ads})} + 2\mathrm{O}^-_{(\mathrm{ads})}. \tag{3.36}$$

The nitrogen dioxide gets ionosorbed onto the surface and therefore occupies a chemisorption site and extracts a conduction electron from the semiconductor. The $O^-_{2(\mathrm{ads})}$ surface species on the other hand disintegrates into two atomic ionized oxygen species, while consuming an

conduction electron. This chemical reaction would lead to

$$\frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} = k_{\mathrm{NO}_2(3)} n_s^2 [\mathrm{NO}_2] N_{\mathrm{O}_2^-} [S]^2.$$

In combination with the equations pertaining to oxygen adsorption (in atomic and molecular form, but both only singly ionized, i.e., $\alpha = 1$) we get the kinetic equations to reaction (3.36), which reads like

$$\frac{dN_{\mathrm{O}_2}}{dt} = k_{\mathrm{O}_2}[\mathrm{O}_2][S] - k_{-\mathrm{O}_2} N_{\mathrm{O}_2} - \frac{dN_{\mathrm{O}_2^-}}{dt}, \tag{3.37a}$$

$$\frac{dN_{\mathrm{O}_2^-}}{dt} = k_{\mathrm{O}_2^-} n_s N_{\mathrm{O}_2} - k_{-\mathrm{O}_2^-} N_{\mathrm{O}_2^-}, - \frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} \tag{3.37b}$$

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[\mathrm{O}_2]^{1/2}[S] - k_{-\mathrm{O}} N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.37c}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-} n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-} N_{\mathrm{O}^-}, + \left(\frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} \right)^{1/2} \tag{3.37d}$$

$$\frac{d[\mathrm{NO}^-_{2(\mathrm{ads})}]}{dt} = k_{\mathrm{NO}_2(3)} [\mathrm{NO}_2] n_s^2 N_{\mathrm{O}_2^-} [S]^2 \tag{3.37e}$$

for $[S] = ([S_t] - N_{\mathrm{O}_2} - N_{\mathrm{O}_2^-} - N_{\mathrm{O}} - N_{\mathrm{O}^-} - [\mathrm{NO}^-_{2(\mathrm{ads})}])$.

It is not clear if the disintegration of adsorbed molecular oxygen is in any way propelled by the adsorption of $\mathrm{NO}_{2(\mathrm{ads})}$, or if it takes place uninfluenced, as is expected in a temperature region around 150 to 250°C, where the dominance of oxygen species on a surface shifts from molecular to atomic forms.

[Ionescu et al. 2003] suggested a reverse reaction instead of the disintegration of $\mathrm{NO}_{2(\mathrm{ads})}$ in equations (3.34) for temperatures above 240°C. As we have done for oxygen we split the ionosorption reaction into a chemisorption and ionization part:

$$\text{Chemisorption:} \qquad \mathrm{NO}_{2(\mathrm{gas})} + S \underset{k_{-\mathrm{NO}_2(3)}}{\overset{k_{\mathrm{NO}_2(3)}}{\rightleftharpoons}} \mathrm{NO}_{2(\mathrm{ads})}, \tag{3.38a}$$

$$\text{Ionization:} \qquad \mathrm{NO}_{2(\mathrm{ads})} + e^- \underset{k_{-\mathrm{NO}_2(4)}}{\overset{k_{\mathrm{NO}_2(4)}}{\rightleftharpoons}} \mathrm{NO}^-_{2(\mathrm{ads})}. \tag{3.38b}$$

[Ionescu et al. 2003] also based their model for oxidizing gases on these reaction equations, although they used a combined reaction mechanism for chemisorption and ionization of nitrogen dioxide. In combination with the reaction equations for oxygen adsorption this would lead to

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[\mathrm{O}_2]^{1/2}[S] - k_{-\mathrm{O}} N_{\mathrm{O}} - \frac{dN_{\mathrm{O}^-}}{dt}, \tag{3.39a}$$

$$\frac{dN_{\mathrm{O}^-}}{dt} = k_{\mathrm{O}^-} n_s N_{\mathrm{O}} - k_{-\mathrm{O}^-} N_{\mathrm{O}^-}, \tag{3.39b}$$

$$\frac{dN_{\mathrm{NO_2}}}{dt} = k_{\mathrm{NO_2}(3)}[\mathrm{NO_2}][S] - k_{-\mathrm{NO_2}(3)}N_{\mathrm{NO_2}} - \frac{dN_{\mathrm{NO_2^-}}}{dt}, \tag{3.39c}$$

$$\frac{dN_{\mathrm{NO_2^-}}}{dt} = k_{\mathrm{NO_2}(4)}n_s N_{\mathrm{NO_2}} - k_{-\mathrm{NO_2}(4)}N_{\mathrm{NO_2^-}}, \tag{3.39d}$$

with $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O^-}} - N_{\mathrm{NO_2}} - N_{\mathrm{NO_2^-}})$, while $N_{\mathrm{NO_2}}$ and $N_{\mathrm{NO_2^-}}$ denote the concentration of nitrogen dioxide adsorbed on the sensor surface in neutral and charged form.

At higher temperatures well above 250°C the copious amounts of $\mathrm{O^-}$ and the increasing amount of trapped surface charge makes a sensor response similar to reducing gases possible, [Ruhland et al. 1998]. This is because the potential barrier at the surface is raised along with the surface charge, making an direct ionosorption of $\mathrm{NO_2}$ to the $\mathrm{SnO_2}$ surface unlikely. Which results in

$$\mathrm{NO_{2(gas)}} + \mathrm{O^-_{(ads)}} \xrightarrow{k_{\mathrm{NO_2}}(5)} \mathrm{NO_{(gas)}} + \mathrm{O_{2(gas)}} + S + e^-, \tag{3.40}$$

where $\mathrm{NO_{2(gas)}}$ participated in an oxidizing reaction with $\mathrm{O^-_{(ads)}}$. The corresponding reaction kinetics in combination with the rate equations for oxygen adsorption are therefore

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[\mathrm{O_2}]^{1/2}[S] - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O^-}}}{dt}, \tag{3.41a}$$

$$\frac{dN_{\mathrm{O^-}}}{dt} = k_{\mathrm{O^-}}n_s N_{\mathrm{O}} - k_{-\mathrm{O^-}}N_{\mathrm{O^-}} - \frac{d[\mathrm{NO_{2(gas)}}]}{dt}, \tag{3.41b}$$

$$\frac{d[\mathrm{NO_{2(gas)}}]}{dt} = k_{\mathrm{NO_2}(5)}[\mathrm{NO_2}]N_{\mathrm{O^-}}, \tag{3.41c}$$

while $[S] = ([S_t] - N_{\mathrm{O}} - N_{\mathrm{O^-}})$.

As the above reaction is likely to deplete the concentration of $\mathrm{O^-}$ on the surface, [Ruhland et al. 1998] suggests a reaction mechanism like (3.34). This leads to the reaction kinetics

$$\frac{dN_{\mathrm{O}}}{dt} = k_{\mathrm{O}}[S][\mathrm{O_2}]^{1/2} - k_{-\mathrm{O}}N_{\mathrm{O}} - \frac{dN_{\mathrm{O^-}}}{dt}, \tag{3.42a}$$

$$\frac{dN_{\mathrm{O^-}}}{dt} = k_{\mathrm{O^-}}n_s N_{\mathrm{O}} - k_{-\mathrm{O^-}}N_{\mathrm{O^-}} + k_{\mathrm{NO_2}(2)}[\mathrm{NO^-_{2(ads)}}], \tag{3.42b}$$

$$\frac{d[\mathrm{NO^-_{2(ads)}}]}{dt} = k_{\mathrm{NO_2}(1)}[\mathrm{NO_2}]n_s[S] - k_{\mathrm{NO_2}(2)}[\mathrm{NO^-_{2(ads)}}], \tag{3.42c}$$

when combined with rate equations for oxygen adsorption (3.5), with $[S] = [S_t] - N_{\mathrm{O}} - N_{\mathrm{O^-}} - [\mathrm{NO^-_{2(ads)}}]$.

Ozone ($\mathrm{O_3}$)

Although ozone is very important for the earth environment as it acts as a protective film in the higher atmosphere, it is also an air pollutant in lower levels of the atmosphere and has

harmful effects on the respiratory system of humans and animals.

In an ozone atmosphere the possible SnO_2 surface reactions are

$$O_{3(\text{gas})} + S + e^- \xrightarrow{k_{O_3}(1)} O^-_{3(\text{ads})}, \tag{3.43a}$$

$$O_{3(\text{gas})} + S + e^- \xrightarrow{k_{O_3}(2)} O_{2(\text{gas})} + O^-_{(\text{ads})}. \tag{3.43b}$$

The first reaction is likely to occur in temperatures well below 150°C, as suggested by [Gurlo et al. 1998], while reaction (3.43b) happens in elevated temperatures above 150°C, according to [Naydenov et al. 1995]. This later reaction is attributed to the unstable character of ozone in elevated temperatures, thus leading to the dissociative reaction in equation (3.43b).

If we once again include this two reactions into the reaction kinetics of oxygen adsorption (in molecular and atomic form, depending on the temperature) we obtain

$$\frac{dN_{O_2}}{dt} = k_{O_2} p_{O_2}[S] - k_{-O_2} N_{O_2} - \frac{dN_{O_2^-}}{dt} \tag{3.44a}$$

$$\frac{dN_{O_2^-}}{dt} = k_{O_2^-} n_s N_{O_2} - k_{-O_2^-} N_{O_2^-} \tag{3.44b}$$

$$\frac{[O^-_{3(\text{ads})}]}{dt} = k_{O_3(1)} p_{O_3} n_s [S] \tag{3.44c}$$

$$[S] = [S_t] - N_{O_2} - N_{O_2^-} - [O^-_{3(\text{ads})}] \tag{3.44d}$$

as probable reaction kinetics for temperatures below 150°C. The reaction kinetics for more elevated temperatures (above 150°C) are given by

$$\frac{dN_O}{dt} = k_O p_{O_2}^{1/2}[S] - k_{-O} N_O - \frac{dN_{O^-}}{dt}, \tag{3.45a}$$

$$\frac{dN_{O^-}}{dt} = k_{O^-} n_s N_O - k_{-O^-} N_{O^-} + \frac{[O_{3(\text{gas})}]}{dt}, \tag{3.45b}$$

$$\frac{[O_{3(\text{gas})}]}{dt} = k_{O_3(2)} p_{O_3} n_s [S] \tag{3.45c}$$

$$[S] = [S_t] - N_O - N_{O^-} \tag{3.45d}$$

Chapter 4

The Combination of different types of Surface Reaction Models and their Solution

4.1 The Interlinking of Surface Reaction Models

In the previous chapter we derived a multitude of surface reaction models, while distinguishing between intrinsic and extrinsic reactions. As intrinsic surface states arise from the dynamics of the sensor alone they should be taken into account for all kind of measurements. Therefore, for measurements in non-inert gases, it is necessary to combine the intrinsic and extrinsic models to form a system of ordinary differential equations. The measurements as well as the estimation of model parameters have to be carried successively from the intrinsic surface state model to the surface reaction model for oxygen to mixtures of synthetic air and one test gas. The different types of surface reaction models therefore form a hierarchy of models.

4.1.1 Intrinsic Model: Sensor response to an inert gas

To explain the response of a sensor in an atmosphere consisting of an inert gas the intrinsic surface state model

$$\frac{dN_{si}}{dt} = k_{\mathrm{i}} n_s (N_i - N_{si}) - k_{-\mathrm{i}} N_{si} \tag{4.1}$$

was adopted. As the concentration of free electrons on the surface can be derived from the concentration of all free electrons (i.e., donors) N_D in the sensor and the height of the potential barrier at the surface V_S, the concentration of free surface electrons n_s is given by (2.2)

$$n_s = N_D \exp\left(-\frac{q^2 N_s^2}{2\varepsilon N_D kT}\right). \tag{4.2}$$

This equation can therefore easily be used to modify equation (4.1), to express n_s as a function of the occupied surface states N_s.

As the reaction rate constants vary with different temperatures, we assume that these reaction rates are of a Arrhenius form. The empirical Arrhenius equation

$$k_x = A_x \exp\left(-\frac{E_x}{kT}\right) \tag{4.3}$$

describes the dependence of rate constants k_x, stemming of chemical reactions x, on the temperature T and the activation energy E_x for reaction x. A_x denotes the pre-exponential constant. The temperature-dependence of A_x can be neglected when compared to the exponential dependence in the second factor. We apply equations (4.2) and (4.3) to the intrinsic surface state model (4.1) and obtain

$$\begin{aligned}\frac{dN_{si}}{dt} = {} & A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{q^2 N_{si}^2}{2\varepsilon N_D kT}\right)(N_i - N_{si}) \\ & - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right) N_{si},\end{aligned} \tag{4.4}$$

which describes the response of a gas sensor to an inert atmosphere.

Independent of subsequent mechanisms for charge transport, this equation is comprised of seven parameters, which need to be found—be it through a parameter estimation procedure or by setting them to reasonable values found in the literature. These seven parameters are: the density of ionized donors N_D, the density of total intrinsic surface states N_i, the pre-exponential constants $A_{\pm\mathrm{i}}$ and the activation energies $E_{\pm\mathrm{i}}$.

4.1.2 Oxygen Model: Sensor Response to dry synthetic air

When describing the response of a sensor to an atmosphere of synthetic air (SA), additionally to the rate equations (3.5) for oxygen adsorption, the intrinsic surface state mechanism has to be taken into account. In this case the behavior of the sensor is due to the superimposition of the intrinsic and extrinsic dynamics. Therefore, after applying Arrhenius equation (4.3) and the formula for free surface electrons (4.2), the surface reaction model describing the sensor response to air is given by

$$\begin{aligned}\frac{dN_{si}}{dt} = {} & A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{q^2 N_s^2}{2\varepsilon N_D kT}\right)(N_i - N_{si}) \\ & - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right) N_{si},\end{aligned} \tag{4.5a}$$

$$\begin{aligned}\frac{dN_{O_\beta^\circ}}{dt} &= A_{O_\beta}\exp\left(-\frac{E_{O_\beta}}{kT}\right)([S_t]-N_{O_\beta^\circ}-N_{O_\beta^{-\alpha}})[O_2]^{\beta/2} \\ &\quad - A_{-O_\beta}\exp\left(-\frac{E_{-O_\beta}}{kT}\right)N_{O_\beta^\circ}-\frac{dN_{O_\beta^{-\alpha}}}{dt},\end{aligned} \tag{4.5b}$$

$$\begin{aligned}\frac{dN_{O_\beta^{-\alpha}}}{dt} &= A_{O_\beta^{-\alpha}}\exp\left(-\frac{E_{O_\beta^{-\alpha}}}{kT}\right)N_D\exp\left(-\frac{q^2N_s^2}{2\varepsilon N_D kT}\right)^\alpha N_{O_\beta^\circ} \\ &\quad - A_{-O_\beta^{-\alpha}}\exp\left(-\frac{E_{-O_\beta^{-\alpha}}}{kT}\right)N_{O_\beta^{-\alpha}},\end{aligned} \tag{4.5c}$$

$$N_s = N_{si}+N_{O_\beta^{-\alpha}}. \tag{4.5d}$$

In this case, the density of occupied surface states N_s is the sum of all occupied surface state—intrinsic and extrinsic. Therefore we can utilize relationship (4.5d) to express the height of the potential barrier due to charged particles trapped at the surface as

$$V_S = \frac{q\left(N_{si}+N_{O_\beta^{-\alpha}}\right)^2}{2\varepsilon N_D}. \tag{4.6}$$

By using relationship (4.2), the concentration of electrons n_s able to reach the surface, depends both on the intrinsic and extrinsic behavior of the sensor, and therefore couples the two dynamics.

Additional to the parameters for the intrinsic model, the following parameters have to be estimated: the concentration of all chemisorption sites on the surface $[S]$, the two pre-exponential factors and two activation energies for the reaction constants of the respective oxygen chemisorption and ionization.

As the type of oxygen species adsorbed at the sensor surface acutely depends on the ambient temperature, the parameters α and β have to be set to the appropriate numbers, depending on the temperature. Below 150°C singly ionized molecular oxygen is dominant, therefore α and β should be set to 1 and 2, respectively. Above this temperature $\alpha = 1$ and $\beta = 1$ would be appropriate. Above 300 to 400°C the singly ionized atomic oxygen species usually gain an additional electron, which means $\alpha = 2$ and $\beta = 1$. As the usual working temperature of tin oxide gas sensors tend to be between 100° and 400°C the setting of $\alpha = 1$ and $\beta = 1$ would be appropriate.

4.1.3 Mixtures of oxygen and an additional gas

In order to model the sensor response to a reducing gas, like CO or an oxidizing agent like NO_2, in addition to synthetic air, a three step approach is taken. As a first and second step

the model parameters and reaction constants for the intrinsic model and the oxygen surface state model are estimated. In a third step the previously gained parameters were incorporated into a whole equation model consisting of the reaction kinetics for the intrinsic surface states and the oxygen surface states (already combined in equations (4.5)) in combination with the model equations for the particular test gas. This necessitates that measurements are taken in an inert atmosphere, in synthetic air as well as in a oxygen and test gas mixture under the same temperature conditions for the same sensor.

In the case of carbon monoxide, the model equations (3.13) are added to (4.4) through the relation of occupied surface states $N_s = N_{si} + N_{O^-} - N_{CO^+}$. As the equations (3.11) regard the direct impact of CO on a gas sensor without an intermediate interaction with oxygen, additional measurements of carbon monoxide in an inert atmosphere are needed to estimate the corresponding model parameters.

For nitrogen dioxide there are a multitude of possible reactions depending on the temperature. Also in this case the selected model equation (3.35), (3.37), (3.39), (3.41) or (3.42) can be linked to the inert model (4.4) through the concentration of occupied surface states N_s.

4.2 Existence and Uniqueness of the Solution

In this section we will investigate the properties of the solution of the surface state models consisting of ordinary differential equations. To accomplish this we will use the Theorem of Picard-Lindelöf, in the usual and generalized version, which can be found in [Cronin 2008] and [Kolgmogorov and Fomin 1970] respectively.

Theorem 4.1 (Picard-Lindelöf). *Let (t_0, x_0) be in an open set D and let a, b be such that the set $R = \{(t,x)|\,|t-t_0| \leq a, |x-x_0| \leq b\}$ is constrained in D. Suppose the function $f(t,x)$ is defined and continuous on D and satisfies a Lipschitz condition with respect to x on R. Then the IVP $x' = f(t,x)$, $x(t_0) = x_0$ has a unique solution on $(t_0 - \min(a, \frac{1}{b}\max_{(t,x)\in R}|f(t,x)|), t_0 + \min(a, \frac{1}{b}\max_{(t,x)\in R}|f(t,x)|))$.*

Proof A proof of this theorem can be found in [Cronin 2008]. □

Theorem 4.2 (Generalization of Picard-Lindelöf). *Let U be an open subset of $\mathbb{R}^{n+1}$ and let the continuous function f be defined as $f(t, x_1, \cdot, x_n) \colon U \subset \mathbb{R}^{n+1} \to \mathbb{R}^n$. If $(t_0, x_{10}, \cdots, x_{n0}) \in U$ and f satisfies the Lipschitz condition*

$$|f_i(t, x_1, \cdots, x_n) - f_i(t, \tilde{x}_1, \cdots, \tilde{x}_n)| \leq L \max_{1\leq j\leq n} |x_j - \tilde{x}_j|$$

in the variables $x_1, \cdots, x_n \in U$*, with* L *a constant. Then the system of ordinary equations*

$$\begin{aligned}\frac{dx_1}{dt} &= f_1(t, x_1, \cdots, x_n),\\ &\vdots\\ \frac{dx_n}{dt} &= f_n(t, x_1, \cdots, x_n)\end{aligned}$$

with initial conditions $x_1(t_0) = x_{10}, \cdots, x_n(t_0) = x_{n0}$ *has an unique solution.*

Proof A proof of this theorem can be found in [Kolgmogorov and Fomin 1970]. □

Based on these results we are able to proof the two important properties of existence and uniqueness of the solution of the surface state models. To be able to more easily handle the surface state model we decided to perform an normalization on the surface state densities N_{si}, $N_{\mathrm{O}^{\circ}_{\beta}}$ and $N_{\mathrm{O}^{-\alpha}_{\beta}}$ to gain the normalized surface state densities nN_{si}, $nN_{\mathrm{O}^{\circ}_{\beta}}$ and $nN_{\mathrm{O}^{-\alpha}_{\beta}}$, which was accomplished by using a normalization factor

$$\alpha = \frac{q}{\sqrt{2\varepsilon N_D k}} \tag{4.7}$$

to transform a densities X into its normalized form $nX = \alpha X = \frac{q\,X}{\sqrt{2\varepsilon N_D k}}$.

4.2.1 Intrinsic Model

In the following Lemma we will prove the existence as well as the uniqueness of the solution to the inert surface state model (4.4) with the help of the Theorem of Picard-Lindelöf.

Lemma 1. *Let a function* $f\colon I \subset \mathbb{R} \to \mathbb{R}$ *, defined as*

$$f(nN_{si}) = A_{\mathrm{i}} \exp\left(\frac{-E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{nN_{si}^2}{T}\right)(nN_i - nN_{si}) - A_{-\mathrm{i}} \exp\left(\frac{-E_{-\mathrm{i}}}{kT}\right) nN_{si}$$

for nonnegative parameters $A_{\pm\mathrm{i}}$*,* $E_{\pm\mathrm{i}}$ *and* nN_i*, be* C^1 *on* D*. When* f *builds the right hand side of the ordinary differential equation* $\frac{d(nN_{si})}{dt} = f(nN_{si})$*, then, for an initial value* $nN_{si}(0) = nN_{si}^0$*, the above differential equation possesses a unique solution.*

Proof To derive an estimated upper bound for the function f we will first calculate the

derivative with respect to nN_{si}

$$f'(nN_{si}) = A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{nN_{si}^2}{T}\right)(1 - nN_i)\,\frac{2nN_{si}}{T}$$
$$- A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{nN_{si}^2}{T}\right) - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right).$$

We use the maximum norm $\|nN_{si}\|_\infty$ to get an upper bound for nN_{si} and can therefore approximate the derivative of f as follows

$$\left|f'(nN_{si})\right| \leq \left|A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D - A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D nN_i\right| \left|\frac{2\|nN_{si}\|_\infty}{T}\right|$$
$$- \left|A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right)\right| = L$$

from the approximation

$$\exp\left(-\frac{nN_{si}^2}{T}\right) \leq 1$$

as nN_{si}^2 is significantly higher than the temperature of 700° K at most. This assumption is justified as values for nN_{si} found in the literature are well above 60. Through the application of the mean value theory we get for some fixed value $\check{n}N_{si}$ in the

$$f'(\check{n}N_{si}) = \frac{f(nN_{si}) - f(n\tilde{N}_{si})}{nN_{si} - n\tilde{N}_{si}} \quad \Rightarrow \quad \left|f(nN_{si}) - f(n\tilde{N}_{si})\right| \leq L \left|nN_{si} - n\tilde{N}_{si}\right|$$

for some constant L. We have therefore shown that f satisfies a Lipschitz condition on I, which allows for the application of the Theorem of Picard-Lindelöf and prove the existence and uniqueness of the solution to the initial value problem $\frac{d(nN_{si})}{dt} = f(nN_{si})$. □

4.2.2 Oxygen Model

To prove the desired properties for the solution of the oxygen model, we can fall back on the proof for the inert model, as the oxygen model incorporates the equation for inert surface states as a third equation. The overall course of action for the oxygen model will essentially be the same as for the intrinsic model, although we make use of the generalization of the Picard-Lindelöf Theorem.

Lemma 2. *Let a function $f\colon D \subset \mathbb{R}^4 \to \mathbb{R}^3$, defined as*

$$\begin{aligned}
f_1(nN_{si}) &= A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{(nN_{si} + nN_{\mathrm{O}_\beta^{-\alpha}})^2}{T}\right)(nN_i - nN_{si}) \\
&\quad - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right) nN_{si}, \\
f_2(nN_{\mathrm{O}_\beta^\circ}) &= A_{\mathrm{O}_\beta} \exp\left(-\frac{E_{\mathrm{O}_\beta}}{kT}\right)([S_t] - nN_{\mathrm{O}_\beta^\circ} - nN_{\mathrm{O}_\beta^{-\alpha}})[\mathrm{O}_2]^{\beta/2} \\
&\quad - A_{-\mathrm{O}_\beta} \exp\left(-\frac{E_{-\mathrm{O}_\beta}}{kT}\right) nN_{\mathrm{O}_\beta^\circ} - \frac{dnN_{\mathrm{O}_\beta^{-\alpha}}}{dt}, \\
f_3(nN_{\mathrm{O}_\beta^{-\alpha}}) &= A_{\mathrm{O}_\beta^{-\alpha}} \exp\left(-\frac{E_{\mathrm{O}_\beta^{-\alpha}}}{kT}\right) N_D \exp\left(-\frac{(nN_{si} + nN_{\mathrm{O}_\beta^{-\alpha}})^2}{T}\right)^{\alpha} nN_{\mathrm{O}_\beta^\circ} \\
&\quad - A_{-\mathrm{O}_\beta^{-\alpha}} \exp\left(-\frac{E_{-\mathrm{O}_\beta^{-\alpha}}}{kT}\right) nN_{\mathrm{O}_\beta^{-\alpha}},
\end{aligned}$$

for nonnegative parameters $A_{\pm x}$, $E_{\pm x}$ and nN_i, be C^1 on D. Let the function f build the right hand side of a system of ordinary differential equations

$$\begin{aligned}
\frac{d(nN_{si})}{dt} &= f_1(nN_{si}), \\
\frac{d(nN_{\mathrm{O}_\beta^\circ})}{dt} &= f_2(nN_{\mathrm{O}_\beta^\circ}), \\
\frac{d(nN_{\mathrm{O}_\beta^{-\alpha}})}{dt} &= f_3(nN_{\mathrm{O}_\beta^{-\alpha}}).
\end{aligned} \tag{4.8}$$

Then, for some initial values, the above differential equation possesses a unique solution.

Proof First we define the vector $\boldsymbol{nN} \in D$ as $(nN_{si}, nN_{\mathrm{O}_\beta^\circ}, nN_{\mathrm{O}_\beta^{-\alpha}})^\top$. After calculating $\nabla_{\boldsymbol{nN}} f_i = (\frac{\partial f_i}{\partial nN_{si}}, \frac{\partial f_i}{\partial nN_{\mathrm{O}_\beta^\circ}}, \frac{\partial f_i}{\partial nN_{\mathrm{O}_\beta^{-\alpha}}})^\top$ for $i = 1, 2, 3$ we use the same approach as in the previous proof to gain an upper bound L_i^j for every single partial derivative $\frac{\partial f_i}{\partial nN_j}$ for $j = si$, O_β°, $\mathrm{O}_\beta^{-\alpha}$. This leads to

$$|\nabla_{\boldsymbol{nN}} f_i(nN_{si}, nN_{\mathrm{O}_\beta^\circ}, nN_{\mathrm{O}_\beta^{-\alpha}})| \leq \boldsymbol{L}_i,$$

for $\boldsymbol{L}_i = (L_i^{si}, L_i^{\mathrm{O}_\beta^\circ}, L_i^{\mathrm{O}_\beta^{-\alpha}})^\top$. We can therefore use the mean value theorem for real valued functions in several variables to gain

$$f_i(\boldsymbol{nN}) - f_i(\tilde{\boldsymbol{nN}}) = \nabla_{\boldsymbol{nN}} f_i(\check{\boldsymbol{nN}})(\boldsymbol{nN} - \tilde{\boldsymbol{nN}}) \leq \boldsymbol{L}_i(\boldsymbol{nN} - \tilde{\boldsymbol{nN}})$$

for some vector $\boldsymbol{n\check{N}} \in \overline{\boldsymbol{nN}\boldsymbol{n}\tilde{\boldsymbol{N}}}$ while $\boldsymbol{nN}$, $\boldsymbol{n}\tilde{\boldsymbol{N}}$, $\boldsymbol{n}\check{\boldsymbol{N}} \in D$. It follows easily that

$$|f_i(\boldsymbol{nN}) - f_i(\boldsymbol{n}\tilde{\boldsymbol{N}})| \leq |\boldsymbol{L}_i(\boldsymbol{nN} - \boldsymbol{n}\tilde{\boldsymbol{N}})| \leq \max_j L_i^j(\max_j |nN_j - n\tilde{N}_j|)$$

for $i = 1, 2, 3$. Hence we can use the Generalized Theorem of Picard Lindelöf to prove the existence and uniqueness of the solution to the system of ordinary differential equations (4.8).□

Chapter 5

Parameter Estimation for Dynamic Models

In this chapter we will deal with the theory behind the estimation of parameters for the type of models relevant to simulate surface reactions—parameter dependent ordinary differential equations. These mathematical models describe the chemical processes on the surface and thus explain the behavior of the observed data. The unknown model parameters will be obtained by minimizing a suitable objective function, which is a measure of the deviation of the data from the model, i.e., the lack of fit. For a survey of parameter estimation for differential equations the reader is referred to [Bard 1974], while [Engl et al. 2009] gives a more general review of inverse problems in applications.

5.1 Structure and Challenges of Dynamical Models

In general, the model equations can be written in the form

$$\frac{d\boldsymbol{x}(\boldsymbol{t})}{d\boldsymbol{t}} = f(\boldsymbol{x}(\boldsymbol{t}), \boldsymbol{u}; \boldsymbol{\theta}) \quad \text{with} \quad \boldsymbol{x}(\boldsymbol{t}_0) = \boldsymbol{x}_0, \tag{5.1a}$$

$$\boldsymbol{y}(\boldsymbol{t}) = \boldsymbol{g}(\boldsymbol{x}(\boldsymbol{t}); \boldsymbol{\theta}), \tag{5.1b}$$

where the included variables and relations are defined as follows:

- $\boldsymbol{\theta} = (\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(P)})^\top$ is the vector of *unknown parameters*, which are quantities like A_i, E_i or N_i in (4.4), that have to be estimated. In this set of parameters some may appear only in equation (5.1a) and others only in equation (5.1b).
- $\boldsymbol{t}$ vector of *independent variables*—e.g., the time.
- $\boldsymbol{u}$ is a vector of variables which are either precisely known or have been measured, e.g., quantities like the Boltzmann constant k.

$\boldsymbol{x} = (x^{(1)}, x^{(2)}, \ldots, x^{(M)})^\top$ defines a vector of *state variables*, which are functions of $\boldsymbol{t}$ and $\boldsymbol{\theta}$, for example N_s. These functions are determined through the ordinary differential equation (5.1a). In a special case the state variables are variables that can be measured through experiments and are therefore observed variables.

$\boldsymbol{x}_0$ is a vector of initial conditions for the state variables.

$\boldsymbol{f}$ defines the M-dimensional vector function forming the right side of the differential equation of state variables.

$\boldsymbol{y} = (y^{(1)}, y^{(2)}, \ldots, y^{(L)})^\top$ is the set of *observed variables* that are measured experimentally, like the conductance G.

$\boldsymbol{g}$ is a known L-dimensional vector function that relates the state vector $\boldsymbol{x}$ to the output vector of observed variables $\boldsymbol{y}$—e.g., the conductance formula (2.14).

Experiments measure the value of $\boldsymbol{y}$ for given values of $\boldsymbol{t}$, which permits the calculation of the values of the state variables $\boldsymbol{x}$ but usually not allow for the calculation of the values of the derivative of $\boldsymbol{x}$. In this case the model equations cannot be used directly for the estimation of the parameters $\boldsymbol{\theta}$. [Bard 1974] lists ways to overcome this problem:

(1) *Differentiation of Data:* To calculate approximate values of the derivative of $\boldsymbol{x}$ in the model equation by calculating a difference quotient between adjacent values. Then $\frac{x_{n+1}-x_n}{t_{n+1}-t_n}$ can give an approximation for $\frac{dx_n}{dt}$. This method is of limited accuracy, it is difficult to access the errors and it is furthermore only applicable if the separation between data points in not too large.

(2) *Integration of Equations:* Depending on the fact if the model equation can be solved analytically or just numerically, the integration of the differential model is a possible way to gain an equation in the form of

$$\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{t}; \boldsymbol{\theta}).$$

If equation (5.1a) can be solved analytically, $\boldsymbol{h}$ can be explicitly given, while for the case that only numerical integration is possible, $\boldsymbol{h}$ is only implicitly defined.

(3) *Integration of Data:* If it is possible to integrate out all the derivatives in the differential equation one gains integral equations. The data may therefore be integrated numerically to obtain the values for the integrals appearing in the model equation. Like for method (1), the data is required to be dense and it is generally only applicable to a limited number of cases, although numerical integration is in general more accurate than numerical differentiation.

Overall, [Bard 1974] recommends to integrate the model equation when computationally possible, and to fall back on the manipulation of data through integration or differentiation if one only wants to obtain an initial guess.

5.1.1 Computation of the objective function

To estimate the parameters $\boldsymbol{\theta}$ we must be able to compute the objective function for any given value of parameters. This usually works by use of some initial estimates for the parameters to gain the initial values in (5.1a). An integration of the differential equation in (5.1a) is therefore possible and gain the values $\boldsymbol{x}_i$, $i = 1, \ldots, N$ of state variables, for N the number of measurements or data points. Equation (5.1b) is then used to calculate values of $\boldsymbol{y}$, which then determine the residuals $\boldsymbol{e}_i = \hat{\boldsymbol{y}}_i - \boldsymbol{y}$, the difference of measured value $\hat{\boldsymbol{y}}_i$ and model calculated value $\boldsymbol{y}_i$. The residuals are then used to calculate the objective function to facilitate the process of parameter estimation. Each function evaluation is therefore in itself a complex procedure that requires the solution of a set of ordinary differential equations, which may be accomplished by the use of standard methods [Bard 1974].

5.1.2 Optimization Methods

To minimize the objective function, like the nonlinear least squares objective function (5.2), a multitude of optimization algorithms may be used. Essentially all these algorithms adhere to the same working principle. These iterative methods start at an initial value for the parameter and iteratively compute a new value while following a rule of calculation, until a minimum is reached. The algorithms find iteratively a new value $\boldsymbol{\theta}_{(j+1)}$ to the current value $\boldsymbol{\theta}_{(j)}$ by moving it along a specified search direction $\boldsymbol{d}_{(j)}$ for a certain amount $s_{(j)}$:

$$\boldsymbol{\theta}_{(j+1)} = \boldsymbol{\theta}_{(j)} + s_{(j)}\boldsymbol{d}_{(j)}.$$

A stopping rule is usually incorporated to specify when the algorithm has found a sufficiently close enough estimation for $\breve{\boldsymbol{\theta}}$, the minimum of the objective function $Q(\boldsymbol{\theta})$. Different algorithms are characterized by their different choices for $s_{(j)}$ and $\boldsymbol{d}_{(j)}$, although the search direction usually in some form incorporates the gradient of the objective function. We set $\boldsymbol{g}(\boldsymbol{\theta})$ to the gradient $\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})$ of the objective function and denote $\boldsymbol{g}_{(j)} = \boldsymbol{g}(\boldsymbol{\theta}_{(j)})$. Furthermore we define $\boldsymbol{H}_{(j)}$ as the Hessian matrix of $Q(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}_{(j)}$.

Steepest Descent Method. This method, also called gradient descent, calculates the parameter values by iterating into the opposite direction of the gradient:

$$\boldsymbol{\theta}_{(j+1)} = \boldsymbol{\theta}_{(j)} - \frac{\boldsymbol{g}_{(j)}^{\top}\boldsymbol{g}_{(j)}}{\boldsymbol{g}_{(j)}^{\top}\boldsymbol{H}_{(j)}\boldsymbol{g}_{(j)}}\boldsymbol{g}_{(j)}.$$

This method is often very inefficient and is not recommended for applications. The steepest descent algorithm may have problems for objective function that, for example are flat around the minimum, as the algorithm may iterate back and forward in this plane when following the inverse direction of the gradient. This problem can be avoided by using the following algorithm.

Newton's Method. This historic algorithm operates by linearizing the objective function and therefore iteratively using the roots of the gradient as an approximation to the root of the objective function:

$$\boldsymbol{\theta}_{(j+1)} = \boldsymbol{\theta}_{(j)} - \boldsymbol{H}_{(j)}^{-1}\boldsymbol{g}_{(j)}.$$

This Algorithm may lead in a false direction when the Hessian is not positive definite, for example when the objective function is concave at $\boldsymbol{\theta}_{(j)}$.

Gauss-Newton Algorithm. By using Newton's Method as a basis and utilizing a quadratic approximation instead of the objective function in every iteration the Gauss-Newton Algorithm reads like

$$\boldsymbol{\theta}_{(j+1)} = \boldsymbol{\theta}_{(j)} + (\nabla_{\boldsymbol{\theta}_{(j)}}\boldsymbol{y}(\boldsymbol{\theta}_{(j)})^\top \nabla_{\boldsymbol{\theta}_{(j)}}\boldsymbol{y}(\boldsymbol{\theta}_{(j)}))^{-1}\nabla_{\boldsymbol{\theta}_{(j)}}\boldsymbol{y}(\boldsymbol{\theta}_{(j)})(\hat{\boldsymbol{\theta}} - \boldsymbol{y}(\boldsymbol{\theta}_{(j)})).$$

The approximation to the Hessian that is used in this algorithm is guaranteed to be positive definite and this method avoids therefore the computation of second order derivatives.

Levenberg-Marquardt Algorithm. This method combines the Gauss-Newton Algorithm with a regularization technique of the Hessian that insures decreasing objective function values. As to insure the positive definiteness of the Hessian $\boldsymbol{H}_{(j)}$ it is possible to correct the inverse of the Hessian at each step by adding a correction matrix to gain a corrected inverse Hessian $\bar{\boldsymbol{H}}_{(j)}^{-1}$ for further computations, such that

$$\bar{\boldsymbol{H}}_{(j)}^{-1} = \boldsymbol{H}_{(j)}^{-1} + c_{(j)}\boldsymbol{I},$$

for an appropriate $c_{(j)} \in \mathbb{R}^+$.

Quasi-Newton Method. This method is an advancement of Newton's Method, as it calculates the inverse of the Hessian $\boldsymbol{H}_{(j)}^{-1}$ by iteratively adding a symmetric correction matrix $\boldsymbol{C}_{(j)}$, such that

$$\bar{\boldsymbol{H}}_{(j)}^{-1} = \boldsymbol{H}_{(j)}^{-1} + \boldsymbol{C}_{(j)}.$$

Prominent examples of Quasi-Newton Methods are the BFGS and DFP algorithm, which differ in the form of the correction matrix.

A more detailed discussion to the above algorithms, especially regarding their computational restrictions or their suitability to specific problems, can be found in [Björck 1996].

Nevertheless, essentially almost any optimization method may be used to minimize the objective function $Q(\boldsymbol{\theta})$. Although, according to [Bard 1974], the Gauss-Newton Method is the most suitable algorithm to solve dynamic models. This is because, as every function evaluation requires the integration of equations (5.1a), the use of a quadratically convergent method like the Newton-Gauss Algorithm is favorable.

5.2 The Objective Function and Nonlinear Least Squares Estimators

The majority of estimation methods found in Estimation Theory, like *maximum a posteriori probability* or *minimum variance estimation*, require statistical knowledge of all or at least some of the random variables involved in a model. The Maximum Likelihood Estimation for example, demands at least knowledge of the probability density of the observation errors, while assuming the statistical properties of the parameters to be estimated are not known. If however no probabilistic information about any of the variables of the models is known, this extreme case, in terms of Estimation Theory, is handled by finding the least squares fit among the observations.

As the objective function is a suitable measure of the overall deviation of the model calculated values from the measurements, its choice dictates not only the values of the parameters but also their statistical properties, [Englezos and Kalogerakis 2001].

We will therefore express the objective function of a parameter estimation problem, that is to be minimized, in terms of Least Squares Estimators. The Least Squares Technique (LS) is one of the most used estimation technique for various applications and can firstly be divided into linear or Ordinary Least Squares (OLS) and Nonlinear Least Squares (NLS). In this section we will mainly concentrate on the nonlinear variant NLS, as the equation occurring in dynamic models are highly nonlinear. Therefore we will use the works of [Englezos and Kalogerakis 2001; Kuan 2004] and [Davidson and MacKinnon 1993] as guide.

The least squares *objective function*

$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{N} \boldsymbol{e}_i(\boldsymbol{\theta})^\top W_i \boldsymbol{e}_i(\boldsymbol{\theta}) \tag{5.2}$$

for the number of measurements N. It is given by the sum of squares of the residuals $\boldsymbol{e}_i(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. The residual equals the difference between measurements and model calculated values of observable variables $\boldsymbol{y}$ (an exact definition will be given below). Depending on the

choice of weighting matrix W_i there are several different cases of estimators:

(1) *Simple Least Squares (SLS) Estimation.* In this case the sum of squares of errors is minimized without a weighting factor. Therefore equation (5.2) reduces to

$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{N} \boldsymbol{e}_i(\boldsymbol{\theta})^\top \boldsymbol{e}_i(\boldsymbol{\theta}),$$

as the weighting matrix can be set to identity, i.e., $W_i = I$.

(2) *Weighted Least Squares (WLS) Estimation.* If the weighting matrix does not change for different experiments ($W_i = W$ for $i = 1, 2, \ldots, N$), the estimator

$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{N} \boldsymbol{e}_i(\boldsymbol{\theta})^\top W \boldsymbol{e}_i(\boldsymbol{\theta})$$

is referred to as Weighted Least Squares.

(3) *Generalized Least Squares (GLS) Estimation.* When the weighted sum of squares of errors is estimated with different weights in every experiment the LS estimator is called Generalized Least Squares Estimator.

For all the above cases, the choice of weighting matrix W can be facilitated by Maximum Likelihood considerations, as can be seen in [Englezos and Kalogerakis 2001]. We will now have a closer look at the simple nonlinear least squares estimation.

5.2.1 Nonlinear Least Squares Estimation

For simplification we assume the univariate case, regarding the observed variables. We consider y to be a scalar and L to be therefore zero. The N distinct measurements of the output vector $\hat{y}$ are related to the value, calculated through the model equation y, using true parameter values $\hat{\boldsymbol{\theta}}$, in the following way:

$$\epsilon_i = \hat{y}_i - y(\boldsymbol{t}_i; \hat{\boldsymbol{\theta}}) \quad \text{for } i = 1, \ldots, N, \tag{5.3}$$

while the *error term* ϵ_i accounts for the measurement error and model inadequacies. Therefore, the output vector $\hat{y}$ is comprised of the *deterministic* part calculated by the model and the *stochastic* part given by the error term. To quantify the deviation of an individual measurement from the model calculated value we state:

Definition (Residual) *The residual of a Nonlinear Least Squares estimator is defined as*

$$e_i = \hat{y}_i - y(\boldsymbol{t}; \boldsymbol{\theta}),$$

where the model function $y(t, \boldsymbol{\theta})$ is evaluated with the use of the estimated parameter values.

To clarify the distinction between the residual and the error term: While the residual e_i corresponds to the estimated parameter values $\boldsymbol{\theta}$, the error term ϵ_i refers to the true parameter values $\hat{\boldsymbol{\theta}}$.

Given the number of measurements N of y and $\boldsymbol{t}$ we combine all measurements into one vector and adopt the notation

$$\hat{\boldsymbol{y}} = \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) + \boldsymbol{e}(\boldsymbol{\theta}),$$

for $\hat{\boldsymbol{y}} = (\hat{y}_1, \ldots, \hat{y}_N)^\top$, $\boldsymbol{e} = (e_1, \ldots, e_N)^\top$ and $\boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) = (y(\boldsymbol{t_1}; \boldsymbol{\theta}), \ldots, y(\boldsymbol{t_N}; \boldsymbol{\theta}))^\top$.

To estimate the parameter vector $\boldsymbol{\theta}$ we can minimize the *LS objective function*

$$\begin{aligned} Q(\boldsymbol{\theta}) &= \frac{1}{N}(\hat{\boldsymbol{y}} - \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}))^\top (\hat{\boldsymbol{y}} - \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})) \\ &= \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - \boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta}))^2, \end{aligned} \tag{5.4}$$

which is given by the weighted sum of squares of the residuals (i.e., the sum of squares error), with respect to $\boldsymbol{\theta}$.

Our goal is tho find a K-dimensional surface that fits the data $(\hat{\boldsymbol{y}}_i, \boldsymbol{x}_i)$, for $i = 1, \ldots, N$. The first order condition of the NLS optimization problem is a system of nonlinear equations with K unknown parameters. We define $\nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) = (\nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1; \boldsymbol{\theta}), \ldots, \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_N; \boldsymbol{\theta}))$ as the gradient of $\boldsymbol{y}$ and can therefore state the *first order optimality condition*

$$\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}) = -\frac{2}{N} \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})(\boldsymbol{y} - \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) = 0. \tag{OC-1}$$

Additionally we define the *second order optimality condition* as

$$\begin{aligned} \nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}) = &- \frac{2}{N} \nabla^2_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{\theta})(\hat{\boldsymbol{y}} - \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})) \\ &+ \frac{2}{N} \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})^\top \quad \text{is positive definite.} \end{aligned} \tag{OC-2}$$

We now can state the definition of a minimizer to the objective function.

Definition (Least Squares Estimator) *The estimated parameter value $\hat{\boldsymbol{\theta}}$ that minimizes the sum of squares error* (5.2)*, by solving* (OC-1) *and satisfying* (OC-2)*, is defined as the Least Squares Estimator and is denoted as $\breve{\boldsymbol{\theta}}$.*

Theorem 5.1. *Let $\boldsymbol{y}(\boldsymbol{x}; \cdot)$ be twice continuously differentiable in the second argument. Given*

the minimization problem

$$\hat{\boldsymbol{y}} = \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta}) + \boldsymbol{e}(\boldsymbol{\theta})$$

we suppose that for given data $(\hat{y}_i, \boldsymbol{x}_i)$*,* $i = 1, \ldots, N$*, the second order condition* (OC-2) *holds for some interior point of* Θ*. Then there exists a solution that minimizes the NLS objective function* (5.4)*.*

While the second order optimality condition ensures that a minimum of S can be found, it is no guarantee to the uniqueness of the solution. This infers the possibility of the existence of multiple local minima for a given data set in the NLS minimization problem. From the first order condition we saw that

$$\nabla_{\boldsymbol{\theta}} Q(\breve{\boldsymbol{\theta}}) = -\frac{2}{N} \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \breve{\boldsymbol{\theta}}) \boldsymbol{e}(\breve{\boldsymbol{\theta}}) = 0$$

Therefore, the residual vector is orthogonal to the column vectors of $\nabla_{\breve{\boldsymbol{\theta}}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \breve{\boldsymbol{\theta}})^\top$. In a geometric sense $\boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})$ defines a surface on Θ and $\nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})^\top$ defines a K-dimensional linear subspace tangent at the point $\boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \boldsymbol{\theta})$ for any $\boldsymbol{\theta} \in \Theta$. Therefore the vector of observables $\hat{\boldsymbol{y}}$ is orthogonally projected onto the mentioned surface at $\boldsymbol{y}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N; \breve{\boldsymbol{\theta}})$, which means that the residual vector $\boldsymbol{e}$ is orthogonal to the tangent space at this point. As, in contrast to the linear case—there may exist more than one such orthogonal projections, there may be multiple solutions to the Nonlinear Least Squares problem.

5.2.2 Consistency of the Nonlinear Least Squares Estimator

A sequence of estimators $\breve{\theta}$ for a parameter θ_0 is said to be (asymptotically) consistent, if $\breve{\theta}$ converges in probability to θ_0. This implies that the distributions of the estimators become more and more concentrated near the true value of the parameter being estimated. In this case the probability that the estimator $\breve{\theta}$ is arbitrarily close to θ_0 converges to one. The definition of this idea is given by

Definition (Consistency) *A sequence of NLS estimators* $\breve{\theta}_N$ *for* $N \geq 0$ *is called strongly (weakly) consistent for the parameter* θ_0 *if* $\breve{\theta}_N \xrightarrow{a.s.(\mathbb{P})} \theta_0$ *as* $N \to \infty$*.*

We first have to give some basic definitions about statistical convergence and the laws of large numbers to facilitate a proof of the consistency of a NLS estimator.

Definition (Stochastic Convergence) *Let* $(t_n)_{n \geq 0}$ *be a sequence of random variables.*

(i) The sequence t_n *converges in probability towards* t *(*$t_n \xrightarrow{\mathbb{P}} t$*), if for all* $\epsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}(|t_n - t| \geq \epsilon) = 0.$$

(ii) The sequence t_n converges almost surely towards t ($t_n \xrightarrow{a.s.} t$), if

$$\mathbb{P}(\lim_{n\to\infty} t_n = t) = 1.$$

(iii) The sequence t_n converges in distribution towards t ($t_n \xrightarrow{\mathbb{D}} t$), if for F_n and F the cumulative distribution functions of t_n and t, if

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for some $x \in \mathbb{R}$, for which F is continuous.

Convergence in probability ensures that the probability of an unlikely outcome becomes smaller and smaller as n grows. Almost sure convergence implies convergence in probability. Stochastic Convergence also introduces a new concept of boundedness, as seen in the following definition.

Definition (Stochastic Boundedness) *A sequence of random variables $(t_n)_{n\geq 0}$ is bounded in probability, if $t_n \xrightarrow{\mathbb{P}} 0$ or if for every $\epsilon > 0$ there is a $\delta < \infty$ such that $\mathbb{P}(|t_n| \geq \delta) < \epsilon$.*

Next we will briefly discuss the Law of Large Numbers (LLN), which concerns the averaging behavior of random variables. There are various types of Laws of Large Numbers for different types of random variables (details can be found in [Kuan 2004]). We will deal with the most elemental of definitions.

Lemma 3 (Law of Large Numbers). *The Strong (Weak) Law of Large Numbers states that the sample average of the sequence t_n converges almost surely (in probability) to the expected value t, that is $\bar{t}_n \xrightarrow{a.s.(\mathbb{P})} t$.*

To apply the idea of a Law of Large Numbers to not only random variables, but functions of random variables, a Uniform Law of Large Numbers (ULLN) is needed.

Lemma 4 (Uniform Law of Large Numbers). *Let $Q(\boldsymbol{\theta})$ be a function of $\boldsymbol{\theta} \in \Theta$. When the conditions*

(1) *$Q(\boldsymbol{\theta})$ obeys a Strong (Weak) Law of Large Numbers for each $\boldsymbol{\theta} \in \Theta$,*

(2) *for $\boldsymbol{\theta}, \boldsymbol{\theta}^\dagger \in \Theta$:*

  (i) *$|Q(\boldsymbol{\theta} - Q(\boldsymbol{\theta}^\dagger)| \leq C||\boldsymbol{\theta} - \boldsymbol{\theta}^\dagger||$ a.s.,*

  (ii) *for C, a random variable which is bounded almost surely (in probability) and is not dependent on $\boldsymbol{\theta}$,*

are met, the function $Q(\boldsymbol{\theta})$ is said to obey a Strong (Weak) Uniform Law of Large Numbers,

which in turn, is defined as

$$\sup_{\theta\in\Theta} |Q(\boldsymbol{\theta}) - \mathbb{E}(Q(\boldsymbol{\theta}))| \xrightarrow{a.s.(\mathbb{P})} 0.$$

Proof A way to derive the property of an Uniform of Large Numbers for a function obeying the above assumption, is shown in [Kuan 2004]. For details and discussion about the necessity of different assumptions for the strong and weak case of the theorem, we also refer to [Kuan 2004]. □

To show the consistency of the NLS estimator we have to introduce one last inequality.

Lemma 5 (Markov's inequality). *Let t be a random variable with finite p^{th} moment. Then Markov's inequality states that*

$$\mathbb{P}(|t| \geq a) \leq \frac{\mathbb{E}(|t|^p)}{a^p}$$

for $a \in \mathbb{R}^+$.

Theorem 5.2 (Consistency). *Let $\boldsymbol{y}(\boldsymbol{x};\cdot)$ be twice continuously differentiable in the second argument. In addition to the the second order condition* (OC-2) *let us assume the conditions*

[C1] *Let θ be in Θ, a compact and convex set, then*

(i) *the sequences $(\hat{y}_i^2)_{i\geq 0}$, $(\hat{y}_i \boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta}))_{i\geq 0}$ and $(\boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta})^2)_{i\geq 0}$ all obey a weak law of large numbers for each $\theta \in \Theta$,*

(ii) *$\hat{y}_i$, $\boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta})$ and $\nabla_{\boldsymbol{\theta}}\boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta})$ have bounded second moment uniformly in $\boldsymbol{\theta} \in \Theta$,*

[C2] *There exists a unique parameter vector $\boldsymbol{\theta}_0$ such that $\mathbb{E}(Q(\boldsymbol{\theta}))$ attains its unique global minimum at $\boldsymbol{\theta}_0$.*

hold. Then the NLS estimator $\breve{\boldsymbol{\theta}}$ is weakly consistent for $\boldsymbol{\theta}_0$.

For an detailed discussion of the conditions [C1] and [C2] the interested reader is referred to [Kuan 2004].

Proof As the condition [C1](i) assures that all components of the NLS objective function $Q(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i^2 - 2\hat{y}_i\boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta}) + \boldsymbol{y}(\boldsymbol{t}_i;\boldsymbol{\theta})^2)$ obey a weak law of large numbers, the first requirement for an ULLN is satisfied.

As our parameter space Θ is compact and convex we use the mean-value theorem as well as the Cauchy-Schwartz inequality to establish the Lipschitz-like continuity condition

$$|Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^\dagger)| \leq \|\nabla_{\boldsymbol{\theta}} Q(\bar{\boldsymbol{\theta}})\| \|\boldsymbol{\theta} - \boldsymbol{\theta}^\dagger\| \text{ a.s.},$$

with $\boldsymbol{\theta}, \boldsymbol{\theta}^{\dagger} \in \Theta$ and $\bar{\boldsymbol{\theta}}$ their mean value. Therefore we set the Lipschitz-type constant to

$$C = \sup_{\boldsymbol{\theta} \in \Theta} \nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}).$$

Hence we have to prove that $\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}) = -\frac{2}{N} \sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta})(\hat{y}_i - \boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta}))$ is bounded in probability. To this end we use condition [C1](ii) to gain

$$\begin{aligned}
\mathbb{P}(|\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})| \geq a) &\leq \frac{1}{a} \mathbb{E}(|\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})|) \\
&\leq \frac{2}{aN} \sum_{i=1}^{N} (\|\nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta})\| \|\hat{y}_i\| + \|\nabla_{\boldsymbol{\theta}} \boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta})\| \|\boldsymbol{y}(\boldsymbol{t}_i; \boldsymbol{\theta})\| \\
&\leq D,
\end{aligned}$$

for some constant D independent of $\boldsymbol{\theta}$. This relation implies that $\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})$ is bounded in probability by Markov's inequality. Therefore we have shown the second condition for an ULLN.

Condition [C2] requires $\boldsymbol{\theta}_0$ to be a global minimum of $\mathbb{E}(Q(\boldsymbol{\theta}))$. Therefore $Q(\boldsymbol{\theta})$ has a WULLN effect, that is

$$\sup_{\theta \in \Theta} |Q(\boldsymbol{\theta}) - \mathbb{E}(Q(\boldsymbol{\theta}))| \xrightarrow{\mathbb{P}} 0.$$

As $\mathbb{E}(Q(\boldsymbol{\theta}))$ is continuous on the compact set Θ, we set $\boldsymbol{\theta}_0$ as the unique, global minimum of $\mathbb{E}(Q(\boldsymbol{\theta}))$. Furthermore, as $\breve{\boldsymbol{\theta}}$ is the NLS estimator we can write

$$Q(\breve{\boldsymbol{\theta}}) = \inf_{\Theta} Q(\boldsymbol{\theta}).$$

Let T be an open neighborhood of $\boldsymbol{\theta}_0$, then we set

$$\epsilon = \inf_{\boldsymbol{\theta} \in T^C \cap \Theta} \mathbb{E}(Q(\boldsymbol{\theta})) - \mathbb{E}(Q(\boldsymbol{\theta}_0)). \tag{5.5}$$

Let A be the event "$|\mathbb{E}(Q(\boldsymbol{\theta})) - Q(\boldsymbol{\theta})| < \epsilon/2$ for all θ". Then

$$A \Rightarrow Q(\breve{\boldsymbol{\theta}}) > \mathbb{E}(Q(\breve{\boldsymbol{\theta}})) - \epsilon/2 \tag{5.6}$$

and

$$A \Rightarrow \mathbb{E}(Q(\boldsymbol{\theta}_0)) > Q(\boldsymbol{\theta}_0) - \epsilon/2. \tag{5.7}$$

Because of $Q(\breve{\boldsymbol{\theta}}) \leq Q(\boldsymbol{\theta}_0)$, we can derive from (5.6) that

$$A \Rightarrow Q(\boldsymbol{\theta}_0) > \mathbb{E}(Q(\breve{\boldsymbol{\theta}})) - \epsilon/2. \tag{5.8}$$

which in turn implies, with the help of equations (5.7) and (5.8), that

$$\begin{aligned} A &\Rightarrow \mathbb{E}(Q(\boldsymbol{\theta}_0)) > Q(\boldsymbol{\theta}_0) - \epsilon/2 > \mathbb{E}(Q(\breve{\boldsymbol{\theta}})) - \epsilon \\ &\Rightarrow \mathbb{E}(Q(\breve{\boldsymbol{\theta}})) - \mathbb{E}(Q(\boldsymbol{\theta}_0)) < \epsilon = \inf_{\boldsymbol{\theta} \in T^C \cap \Theta} \mathbb{E}(Q(\boldsymbol{\theta})) - \mathbb{E}(Q(\boldsymbol{\theta}_0)) \\ &\Rightarrow \mathbb{E}(Q(\breve{\boldsymbol{\theta}})) < \inf_{\boldsymbol{\theta} \in T^C \cap \Theta} \mathbb{E}(Q(\boldsymbol{\theta})) \\ &\Rightarrow \breve{\boldsymbol{\theta}} \in \Theta. \end{aligned} \tag{5.9}$$

We therefore see that $A \Rightarrow \breve{\boldsymbol{\theta}} \in T$, from which we can conclude that $\mathbb{P}(A) \leq \mathbb{P}(\breve{\boldsymbol{\theta}} \in T)$. And as the WULLN property of $Q(\boldsymbol{\theta})$ implies that $\lim \mathbb{P}(A) = \lim \mathbb{P}(|Q(\boldsymbol{\theta}) - \mathbb{E}(Q(\boldsymbol{\theta}))| < \epsilon/2) = 1$, we see that $\lim \mathbb{P}(\breve{\boldsymbol{\theta}} \in T) = 1$. As B was arbitrary, $\breve{\boldsymbol{\theta}}$ must converge to $\boldsymbol{\theta}_0$ in probability. □

As the above theorem is concerned with the convergence to a global minima instead of a local one, it is not completely satisfactory. This is because an iterative algorithm can, in general, only be expected to find local minima. According to [Kuan 2004], a simple proof for local consistency is not yet available, so they only state that the Nonlinear Least Squares Estimator converges in probability to a local minimum.

5.2.3 Asymptotic Normality

Definition (Asymptotic Normality) *An estimator $\breve{\boldsymbol{\theta}}$ that is consistent for $\boldsymbol{\theta}_0$, is asymptotically normal if the sequence of random variables $(\sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0))_{N \geq 0}$ tends in distribution to a normal distribution with zero mean an finite covariance matrix for growing sample size N.*

To prove asymptotic normality we also need the Central Limit Theorem (CLT). If a CLT holds, the distributions of suitably normalized averages of random variables are close to the standard normal distribution in the limit, regardless of the original distributions of these random variables. As for the Law of Large Numbers there are different CLTs for different kinds of random variables, [Kuan 2004].

Lemma 6 (Central Limit Theorem). *The Central Limit Theorem states for the sample average $\bar{t}_n$ of a sequence of random variables t_n with mean μ and variance $\sigma^2 > 0$ that*

$$\frac{\sqrt{N}(\bar{t}_n - \mu)}{\sigma} \xrightarrow{D} \mathcal{N}(0, 1).$$

Theorem 5.3 (Asymptotic Normality). *Let $\boldsymbol{\theta}$ be consistent for $\boldsymbol{\theta}_0$. In addition to the the first and second order conditions* (OC-1) *and* (OC-2) *let us assume the following conditions hold*

[N1] *The sequence $(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}))_{N\geq 0}$ obeys a WULLN.*

[N2] *$\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}))$ is continuous in $\boldsymbol{\theta}$.*

[N3] *The sequence $(\sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}))_{N\geq 0}$ obeys a CLT.*

Then the NLS estimator $\breve{\boldsymbol{\theta}}$ is asymptotically normal.

Proof We first use the mean value theorem to expand $\nabla_{\boldsymbol{\theta}} Q(\breve{\boldsymbol{\theta}})$ about $\boldsymbol{\theta}_0$ to gain

$$\nabla_{\boldsymbol{\theta}} Q(\breve{\boldsymbol{\theta}}) = \nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0) + \nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0),$$

with $\boldsymbol{\theta}^\dagger$ the mean value of $\breve{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}_0$. As $\breve{\boldsymbol{\theta}}$ is the NLS estimator and solves the first order condition (OC-1), the left hand side must be zero. As the second order condition holds, the Hessian is invertible and therefore

$$\sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)^{-1}\sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0).$$

Which implies that $\sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ and the right hand side are of the same distribution. Let $A^\perp$ denote the vector formed by adjoining the rows of the matrix A to each other. Therefore

$$\begin{aligned} &\|(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)) \perp -(\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0)))^\perp\| \\ &\quad \leq \|(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger))^\perp - (\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)))^\perp\| \\ &\qquad - \|(\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)))^\perp - (\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0)))^\perp\|, \end{aligned} \tag{5.10}$$

according to the triangle equation. Because of $\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})$ obeying a WULLN, the first term in (5.10) converges to zero in probability. The consistency of $\breve{\boldsymbol{\theta}}$ also implies that the mean value $\boldsymbol{\theta}^\dagger$ converges to $\boldsymbol{\theta}_0$ in probability. This, given condition [N2] holds, implies that the second term of relation (5.10) also converges to zero in probability, which implies that $\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}^\dagger)) \xrightarrow{\mathbb{P}} \mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0))$ and that $\sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is of the same distribution as $-\mathbb{E}(\nabla^2_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0))^{-1}\sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0)$.

As, according to [N3], $\sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0)$ obeys a CLT, we have

$$\left(\sqrt{\mathbb{V}ar(\sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0))}\right)^{-1} \sqrt{N}\nabla_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}_0) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{1}),$$

which in turn implies

$$\left(\sqrt{\boldsymbol{V}_0}\right)^{-1} \sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{1})$$

for $\boldsymbol{V}_0$ the covariance matrix (For details regarding the composition of the covariance matrix, see [Kuan 2004; Davidson and MacKinnon 1993]). We have therefore shown that $\sqrt{N}(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ follows a zero mean normal distribution. □

Chapter 6

Quantitative Analysis and Simulation

We will now have a closer look at the behavior of the inert gas model (4.4) in combination with the conductance formulas (2.14) stated in Chapter 2. We will consider the equations in the normalized form, after applying the normalization factor α from equation (4.7), as shown in Chapter 4. The model equations therefore read as

$$\frac{d(nN_{si})}{dt} = A_{\mathrm{i}} \exp\left(-\frac{E_{\mathrm{i}}}{kT}\right) N_D \exp\left(-\frac{nN_{si}^2}{T}\right) (nN_i - nN_{si}) - A_{-\mathrm{i}} \exp\left(-\frac{E_{-\mathrm{i}}}{kT}\right) nN_{si}, \tag{6.1a}$$

$$G\left(T, V_s\right) = \begin{cases} G_0^* T^{-3/2} \left(1 - \beta n N_s\right) \exp\left(-\frac{\gamma}{T}\right) & \text{complete depletion of electrons,} \\ G_0^* T^{-3/2} \exp\left(-\frac{nN_s^2}{T}\right) & \text{partial depletion of electrons,} \end{cases} \tag{6.1b}$$

while some factors in the conductance formula for the completely depleted case where combined into the parameters β and γ, as was recommended by [Fort et al. 2010]. Also the temperature dependence of the pre-exponential factor G_0 was expressed separately.

We decided to not make use of the additional term G_C in the conductance models (2.14) while estimating the model parameters, as the corresponding values for G_C reported in the work of Fort et al. were generally very small.

6.1 Description of Data

The data, which was graciously provided by the Austrian Institute of Technology (AIT), was derived by measuring the change in voltage for a constant current of 1 nA generated by a bundle of nanowires acting as sensing element. The measurements where taken in a dry

nitrogen atmosphere, while the temperature was changed. The corresponding temperature profile can be seen in Figure 6.1.

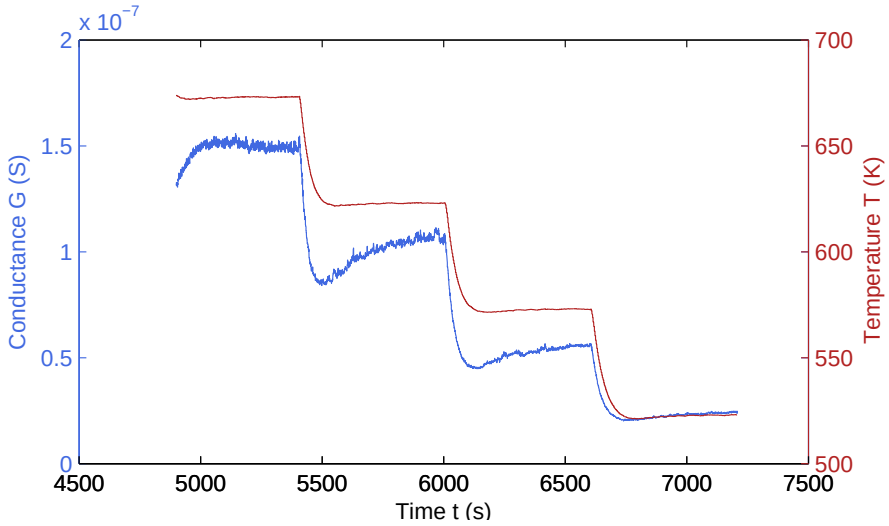


Figure 6.1: Plot of the measured conductance G and the associated temperature profile.

The response of the sensor to the temperature profile, conforms to the intrinsic surface state trapping model of Ding et al., which was discussed in Chapter 3.2.

The steep drop in conductance, as a reaction to a decrease in temperature, is a consequence of the fast dynamic process of electrons dropping from the conduction band into the valence band. As the release of electrons from occupied intrinsic surface states into the conduction band is a slower dynamic process, the conductance slowly increases again after every temperature drop until either thermal equilibrium is reached or another temperature drop occurs.

As the data shows the response of a bundle of nanowires to an inert atmosphere the model equations (6.1) are applicable.

6.2 Calculation of nN_s from measured data

The preliminary calculation of surface states nN_s from the measured sensor conductance is conducted essentially through formula

$$I = \frac{n \cdot q}{t},$$

which states that the current depends on the flow of charge through the conducting channel. I states the current in Ampere, n is the number of charge carriers and t denotes the time (in

seconds) over which the flow occurs. As a constant current is applied to the sensor during measurements, the in-/decrease in the measured conductance is caused by the trapping and releasing of electrons in/from surface states.

As the current was set to a constant 1 nA and a change in conductance was observed during measurements, it is assumed that a part of the current was produced by the change in the surface state density. This part of current is proportional to the measured change in conductance. Thus it is possible to deduce the part of current and therefore the number of electrons stemming from a change in the intrinsic surface states density which caused the measured change in conductance:

$$n_s = \frac{G_{t+1}/G_t \cdot I \cdot t}{q} = \frac{b \cdot 10^{-2} \cdot 10^{-9} \cdot t}{q}$$

with t the time interval over which the change in conductance $G_{t+1}/G_t = b\%$ occurred. As the concentration of free surface electrons n_s is given by bulk electrons N_D energized enough to overcome the surface potential barrier V_s, with the help of equations (2.2) and (2.10), equation

$$n_s = N_D \exp\left(-\frac{nN_s^2}{T}\right) \quad \Rightarrow \quad nN_s = \frac{\sqrt{-T \log\left(\frac{b \cdot t \cdot 10^{-11}}{q\ N_D}\right)}}{q} \tag{6.2}$$

allows us to calculate the number of occupied surface states form the measured conductance.

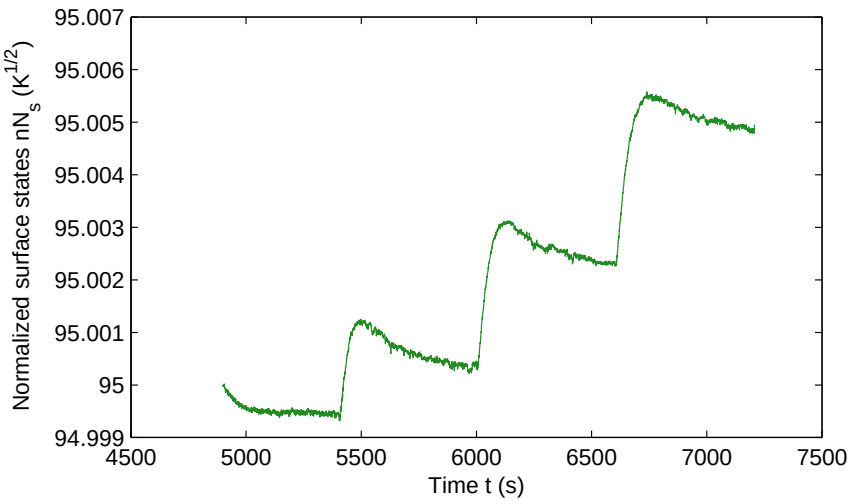


Figure 6.2: A picture showing the form of the normalized surface states calculated through the current flow equation (6.2).

The occupied surface states calculated in that manner, follow changes in the sensor conductance very closely. Which also implies that the nN_s would react to abrupt changes in temperature as swiftly as the conduction does. This indicates a different behavior of nN_s than Ding et al. predicted.

The parameters for the conductance formulas of (6.1b) where estimated by fitting a nonlinear least squares error in the cases of complete and partial depletion. The resulting parameters are shown in Table 6.1.

Case of carrier depletion	Parameters	G_0^*	β	γ
Complete depletion	Values	1.396e+02	1.050e−02	3.306e+03
Partial depletion		1.952e+03		

Table 6.1: Conduction parameter sets for the cases of partial and complete depletion.

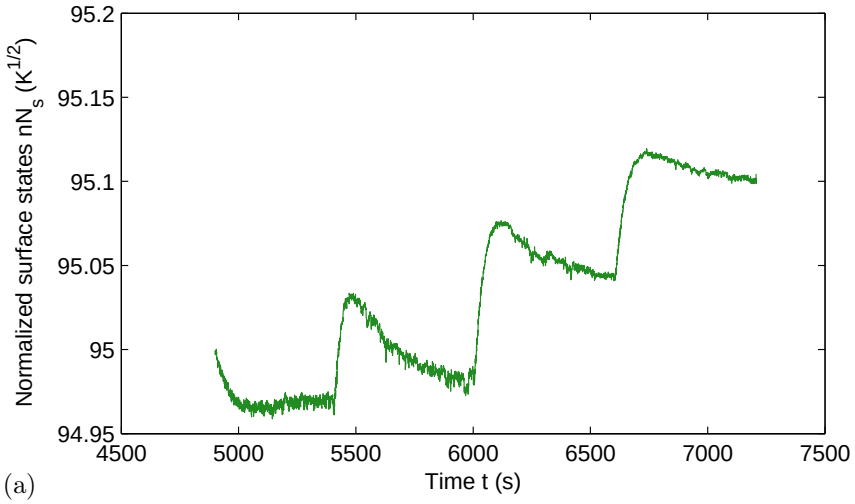
The conductance parameters in Table 6.1 were then used to calculate the normalized surface state densities nN_s for both cases of charge depletion from the measured conductance G.

The first plot in Figure 6.3 refers to the density of normalized surface states calculated under the assumption of a complete depletion of charge carriers throughout the nanowires. In this case the nN_s behave quite similar to the normalized surface states calculated through the current flow formula (6.2). This is due to the fact that the second term in

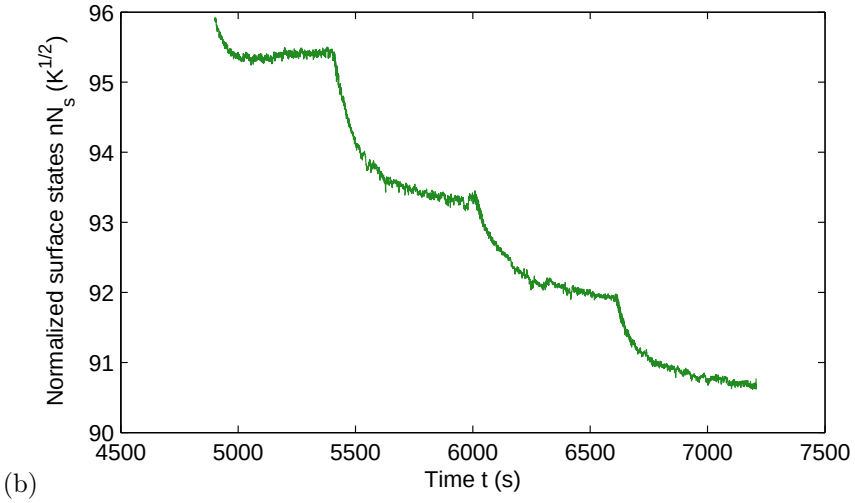
$$nN_s = \frac{1}{\beta}\left(1 - \exp\left(\frac{\gamma}{T}\right)\frac{G}{G_0}T^{3/2}\right), \tag{6.3}$$

which was derived from (6.1b), dictates a form negative to the conductance for the nN_s. This also implies the immediate response of nN_s to changes of temperature, which does not conform to the intrinsic surface state trapping model of Ding et al. If one would consider the approximate real values for the parameters β and γ (which would be of order 10^{-6} and 10^{11} respectively), no plausible values for nN_s would be reached, as the second term in (6.3) would be too high, resulting in extremely high, negative values for the nN_s.

Figure 6.3(b) shows that the surface state density behaves like the intrinsic surface state trapping model predicts. The N_s react slowly to a step in temperature and decrease as the conductance increases.



(a)



(b)

Figure 6.3: Plot of the normalized surface state density nN_s for the completely (a) and partly depleted case (b).

6.3 Simulation of nN_s and G

To estimate the parameters in the surface model (6.1a) we will use a nonlinear least squares objective function as well as the MATLAB function *lsqnonlin*. We will consider the conductance model (6.1b) in the case of partial depletion of charge carriers. Although one would expect for A_i to have a similar value to the reaction rate A_{-i} of its inverse reaction, the fitted parameters show similar values as seen in the work of Fort et al.

Parameters	A_i	E_i	nN_i	A_{-i}	E_{-i}
Values	1.237e−08	1.293e+04	9.593e+01	2.543e+03	1.011e+04

Table 6.2: The fitted parameters for the intrinsic surface state model (6.1a).

On basis of these parameters the normalized surface states nN_s were computed and can be seen in Figure 6.4. They show a somewhat better resemblance to the values calculated from the data for the first two temperature steps than for the third jump. Despite that, the associated fitting error is below 1.5%. This nevertheless implies that a lot of additional work has to be put into the fitting of surface reaction parameters to gain simulated surface states that are in a better agreement to the values of Figure 6.3(b)—especially regarding a more correct simulation of the curvature in the surface states as a reaction to the third temperature step.

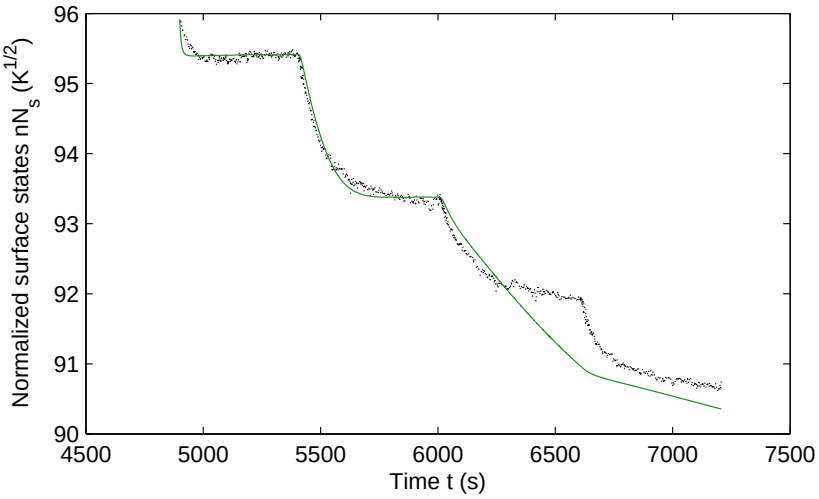


Figure 6.4: Simulated nN_s computed with surface parameter set in Table 6.2 vs. measured values.

The subsequent calculation of the corresponding conductance values are shown in Figure 6.5. The fitting error of the above simulation is now propagated to the simulation of the conductance, leading to an error of 27.9% at the third temperature step, while up until this point it was well below 10%.

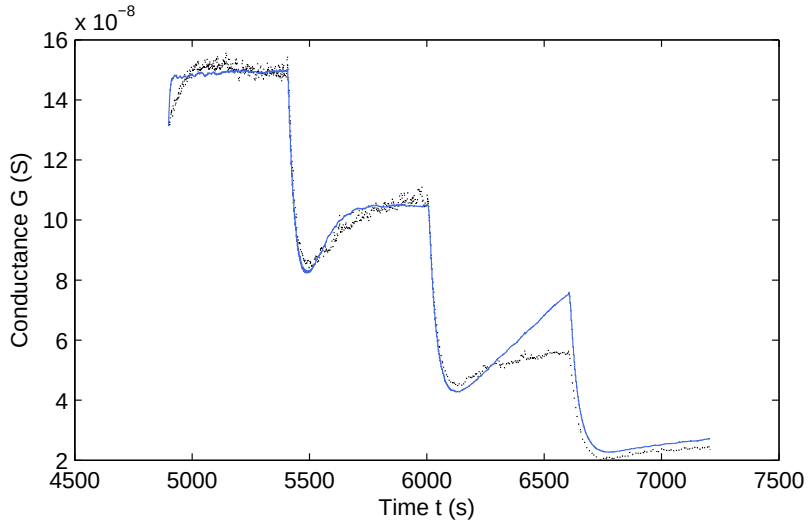


Figure 6.5: Simulated conductance G for computed nN_s values from parameter set in Table 6.2 vs. measured conductance values.

Bibliography

Y. Bard. *Nonlinear Parameter Estimation.* Academic Press, 1974.

N. Bârsan and U. Weimar. Conduction Model of Metal Oxide Gas Sensors. *Journal of Electroceramics*, **7**:143–167, 2001.

N. Bârsan and U. Weimar. Understanding the fundamental principles of metal oxide based gas sensors; the example of CO sensing with SnO_2 sensors in the presence of humidity. *J. Phys.: Condens. Matter*, **15**:R813–R839, 2003.

A. Björck. *Numerical Methods for Least Squares Problems.* SIAM, 1996.

A. Broniatowski. *Polycrystalline Semiconductors*, pages 95–117. Springer, 1985.

Y.-J. Choi, I.-S. Hwang, J.-G. Park, K. J. Choi, J.-H. Park, and J.-H. Lee. Novel fabrication of an SnO_2 nanowire gas sensor with high sensitivity. *Nanotechnology*, **19**, 2008.

E. Comini, C. Baratto, G. Faglia, M. Ferroni, A. Vomiero, and G. Sberveglieri. Quasi-one dimensional metal oxide semiconductors: Preparation, characterization and application as chemical sensors. *Progress in Materials Science*, **54**:1–67, 2009.

J. Cortés, A. Narváez, H. Puschmann, and E. Valencia. Mean field theory studies of surface reactions on disordered substrates. *Chemical Physics*, **288**:77–88, 2003.

J. Cronin. *Ordinary Differential Equations: introduction and qualitative theory.* Chapman & Hall/CRC, 2008.

R. Davidson and J. G. MacKinnon. *Estimation and Inference in Econometrics.* Oxford University Press, 1993.

R. Dickman. Kinetic phase transitions in a surface-reaction model: Mean-field theory. *Physical Review A*, **34**:4246–4250, 1986.

J. Ding, T. J. McAvoy, R. E. Cavicchi, and S. Semancik. Surface state trapping models for SnO_2-based microhotplate sensors. *Sensors and Actuators B*, **77**:597–613, 2001.

H. W. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, and P. Schuster. Inverse problems in systems biology. *Inverse Problems*, **25**:123014, 2009.

P. Englezos and N. Kalogerakis. *Applied Parameter Estimation for Chemical Engineers*. Marcel Dekker, Inc., 2001.

A. Fort, M. Mugnaini, S. Rocchi, M. B. Serrano-Santos, R. Spinicci, and V. Vignoli. Surface State Model for Conductance Responses During Thermal-Modulation of SnO_2-Based Thick Film Sensors: Part II - Experimental Verification. *IEEE Transactions on Instrumentation and Measurement*, **55**:2107–2117, 2006a.

A. Fort, S. Rocchi, M. B. Serrano-Santos, R. Spinicci, and V. Vignoli. Surface State Model for Conductance Responses During Thermal-Modulation of SnO_2-Based Thick Film Sensors: Part I - Model Derivation. *IEEE Transactions on Instrumentation and Measurement*, **55**: 2102–2106, 2006b.

A. Fort, M. Mugnaini, S. Rocchi, M. B. Serrano-Santos, V. Vignoli, and R. Spinicci. Simplified models for SnO_2 sensors during chemical and thermal transients in mixtures of inert, oxidizing and reducing gases. *Sensors and Actuators B*, **124**:245–259, 2007.

A. Fort, M. Mugnaini, V. Vignoli, S. Rocchi, E. Comini, G. Faglia, and A. Ponzoni. Characterization and modelling of SnO_2 nanowire sensors for CO detection. In *Proceedings of the 3rd International Workshop on Advances in sensors and Interfaces*, pages 41–45, 2009.

A. Fort, M. Mugnaini, S. Rocchi, V. Vignoli, E. Comini, G. Faglia, and A. Ponzoni. Metal-oxide nanowire sensors for CO detection: Characterization and modeling. *Sensors and Actuators B*, **148**:283–291, 2010.

D. Girardin, F. Berger, A. Chambaudet, and R. Planade. Modelling of SO_2 detection by tin dioxide gas sensors. *Sensors and Actuators B*, **43**:147–153, 1997.

J. Gong, Q. Chen, W. Fei, and S. Seal. Micromachined nanocrystalline SnO_2 chemical gas sensors for electronic nose. *Sensors and Actuators B*, **102**:117–125, 2004.

A. Gurlo, N. Bârsan, M. Ivanovskaya, U. Weimar, and W. Göpel. In_2O_3 and $MoO_3 - In_2O_3$ thin film semiconductor sensors interaction with NO_2 and O_3. *Sensors and Actuators B*, **47**:92–99, 1998.

S. H. Hahn, N. Bârsan, U. Weimar, S. G. Ejakov, J. H. Visser, and R. E. Soltis. CO sensing with SnO_2 thick film sensors: role of oxygen and water vapor. *Thin Solid Films*, **436**:17–24, 2003.

J. Häusler. *Charakterisierung von Gassensoren zur Überwachung belasteter Raumluft*. PhD thesis, Justus Liebig University Gießen, Germany, 2004.

D. J. Higham. Modeling and Simulating Chemical Reactions. *SIAM Review*, **50**:247–368, 2008.

N. D. Hoa, N. V. Quy, and D. Kim. Nanowire structured SnO_x-SWNT composites: High performance sensor for NO_x detection. *Sensors and Actuators B*, **142**:153–259, 2009.

J. Huang and Q. Wan. Gas sensors based on semiconducting metal oxide one-dimensional nanostructures. *Sensors*, **9**:9903–9924, 2009.

R. Ionescu, E. Llobet, S. Al-Khalifa, J. W. Gardner, X. Vilanova, J. Brezmes, and X. Correig. Response model for thermally modulated tin oxide-based microhotplate gas sensors. *Sensors and Actuators B*, **95**:203–211, 2003.

K. Katterbauer. Pbsens: A mathematical model for nanowire gas-sensors. Master's thesis, University of Vienna, 2010.

U. Kersen and L. Holappa. H_2S-sensing properties of SnO_2 produced by ball milling and different chemical reactions. *Analytica Chimica Acta*, **562**:110–114, 2006.

A. Köck, A. Tischner, T. Maier, M. Kast, C. Edtmaier, C. Gspan, and G. Kothleitner. Atmospheric pressure fabrication of SnO_2-nanowires for highly sensitive CO and CH_4 detection. *Sensors and Actuators B*, **138**:160–167, 2009.

D. Kohl. Function and applications of gas sensors. *J. Phys. D: Appl. Phys.*, **34**:R125–R149, 2001.

A. N. Kolgmogorov and S. V. Fomin. *Introductory Real Analysis*. Dover Publications, Inc., 1970.

C.-M. Kuan. Introduction to econometric theory. Lecture Notes, 2004. URL `http://idv.sinica.edu.tw/ckuan/pdf/et01/et01.pdf`.

G. Leo, R. Rella, P. Siciliano, S. Capone, J. Alonso, V. Pankov, and A. Ortiz. Sprayed SnO_2 thin films for NO_2 sensors. *Sensors and Actuators B*, **58**:370–374, 1999.

H. Liu, S. P. Gong, Y. X. Hu, J. Q. Liu, and D. X. Zhou. Properties and mechanism study of SnO_2 nanocrystals for H_2S thick-film sensors. *Sensors and Actuators B*, **140**:190–195, 2009.

J. G. Lu, P. Chang, and Z. Fan. Quasi-one-dimensional metal oxide materials: Synthesis, properties and applications. *Materials Science and Engineering R*, **52**:49–91, 2006.

E. W. Lund. Guldberg and waage and the law of mass action. *Journal of Chemical Education*, 42:548–550, 1965.

M. J. Madou and S. R. Morrison. *Chemical Sensing With Solid State Devices*. Academic Press, first edition, 1989.

V. V. Malyshev. Response of Semiconducting Metal Oxides to Water Vapor as a Result of Water Molecules Chemical Transformations on Catalytically Active Surface. *Russian Journal of Physical Chemistry A*, **82**:2329–2339, 2008.

V. V. Malyshev and A. V. Pislyakov. SnO_2-based thick-film-resistive sensor for H_2S detection in the concentration range of 1-10 mg m^{-3}. *Sensors and Actuators B*, **47**:181–188, 1998.

V. V. Malyshev and A. V. Pislyakov. Investigation of gas-sensitivity of sensor structures to hydrogen in a wide range of temperature, concentration and humidity of gas medium. *Sensors and Actuators B*, **134**:913–921, 2008.

S. R. Morrison. *The Chemical Physics of Surfaces*. Plenum Press, second edition, 1990.

A. Naydenov, R. Stoyanova, and D. Mehandjiev. Ozone decomposition and CO oxidation on CeO_2. *Journal of Molecular Catalysis A: Chemical*, **98**:9–14, 1995.

H.-Y. Pan and H. J. Wang. A two-species surface reaction model of the mixing type. *Physica A*, **227**:234–238, 1996.

F. Quaranta, R. Rella, P. Siciliano, S. Capone, M. Epifani, L. Vasanelli, A. Licciulli, and A. Zocco. A novel gas sensor based on SnO_2/Os thin film for the detection of methane at low temperature. *Sensors and Actuators B*, **58**:350–355, 1999.

B. Ruhland, T. Becker, and G. Müler. Gas-kinetic interaction of nitrous oxides with SnO_2 surfaces. *Sensors and Actuators B*, **50**:85–94, 1998.

I. Sayago, J. Gutihrez, L. Ads, J. Robla, M. Horrillo, J. Getino, and J. Agapito. The interaction of different oxidizing agents on doped tin oxide. *Sensors and Actuators B*, **24-25**:512–515, 1995.

K. Tabata, T. Kawabe, Y. Yamaguchi, and Y. Nagasawa. Chemisorbed oxygen species over the (110) face of SnO_2. *Catalysis Surveys from Asia*, **7**:251–259, 2003.

W. E. Taylor, N. H. Odell, and H. Y. Fan. Grain boundary barriers in germanium. *Physical Review*, **88**:867–875, 1952.

A. Tischner, T. Maier, C. Stepper, and A. Köck. Ultrathin SnO_2 gas sensors fabricated by spray pyrolysis for the detection of humidity and carbon monoxide. *Sensors and Actuators B*, **134**:796–802, 2008.

A. Tischner, A. Köck, T. Maier, C. Edtmaier, C. Gspan, and G. Kothleitner. Tin oxide nanocrystalline films and nanowires for gas sensing applications. *Microelectronic Engineering*, **86**:1258–1261, 2009.

Z. L. Wang. ZnO nanowire and nanobelt platform for nanotechnology. *Material Science and Engineering R*, **64**:33–71, 2009.

N. Yamazoe and K. Shimanoe. Theory of power laws for semiconductor gas sensors. *Sensors and Actuators B*, **128**:566–573, 2008.

N. Yamazoe and K. Shimanoe. Theoretical approach to the rate of response of semiconductor gas sensors. *Sensors and Actuators B*, **154**:277–282, 2010.

A. Zima. *Development of highly sensitive nano-gassensors based on nanocrystalline tin dioxide thin film and single-crystalline tin dioxide nanowires.* PhD thesis, Vienna University of Technology, 2009.

A. Zima, A. Köck, and T. Maier. In- and sb-doped tin oxide nanocrystalline films for selective gas sensing. *Microelectronic Engineering*, **87**:1467–1470, 2010.

Acknowledgments

I would like to take this opportunity to express sincere gratitude to my supervisor Clemens Heitzinger, for giving me the chance to write my diploma thesis in a fascinating field of science with real life application, for his technical expertise, ideas, time and patience.

I also thank the WPI (Wolfgang Pauli Institute) and its members in general, for providing an ideal setting for my thesis by supplying a inspiring working environment. A special thanks to Klemens Katterbauer, with whom I have discussed many ideas and who laid the computational groundwork. In particular I would like to thank Alena Bulyha, Martin Vasicek, Stefan Baumgartner, Nathalie Tassotti and Angelika Manhart, with whom I have been fortunate enough to benefit from working with. Also I would like to thank Elise Brunet and Anton Köck as well as their colleagues from the AIT for giving explanations to chemical and physical aspects of gas sensors and specifically for supplying the relevant measurement data.

Furthermore I would like to acknowledge my family, and most of all my parents Marina and Georg Rehrl for supporting and encouraging me in my studies and for making them possible in the first place.

I also want to thank all my friends for backing me up in my daily life.

Financial support by the Vienna Science and Technology Fund (WWTF) project "Mathematics and Nanosensors" (No.MA09-028) is thankfully acknowledged.

Abstract

This thesis deals with the modeling of metal oxide nanowire gas sensors, which are technical devices in the scale of nanometers, that are used to detect different gases in different concentrations in the atmosphere. As state of the art gas sensors show a low selectivity, the detailed modeling of the surface reactions caused by different test gases is essential to overcome this issue.

The developed response models, described in this work, are composed of an ODE surface reaction model and a charge transport model (both parameter dependent) and predict the change of conductance of the sensor upon changes in the thermal and chemical environment. In this diploma thesis we present surface reaction models for the most important and test gas species that are relevant for applications, show their derivation and give a detailed discussion of their properties.

In order to facilitate the simulation of the sensor response, the theory of inverse modeling of dynamic models, with special regard to the estimation of model parameters through nonlinear least squares estimators and suitable optimization algorithms, is discussed.

A one parameter transport model in combination with a five parameter surface reaction model is used to simulate the response of a gas sensor consisting of a bundle of nanowires in an inert atmosphere. Reaction as well as conduction parameters is fitted in a nonlinear least squares estimation process, using inverse modeling techniques.

Zusammenfassung

Diese Diplomarbeit behandelt die mathematische Modellierung von Nanowire Gassensoren, also Sensoren in der Größenordnung einiger Nanometer die zur Detektion verschiedener Gase und deren Konzentration in der Umgebungsluft verwendet werden. Aktuell erforschte und enwickelte Gassensoren können schwer zwischen bestimmten Gasen unterscheiden, desswegen ermöglicht die mathematischen Modellierung ein tiefergehendes Verständnis der für den Sensorausschlag verantwortlichen Oberflächereaktionen auch eine Verbesserung dieser Kreuzselektivität.

Die in dieser Arbeit hergeleiteten Sensormodelle setzen sich jeweils aus einem Oberflächenmodell, bestehend aus gewöhlichen Differentialgleichungen, und einem Ladungstransportmodell – beide paramterabhängig – zusammen und sagen die Änderung der Leitfähigkeit, verursacht durch eine Änderung der Temperature und der chemischen Zusammensetzung der Umgebungsluft, voraus. Wir präsentieren Oberflächenmodelle für alle wichtigen, für Anwendungen relevanten Gassorten, leiten sie her und diskutieren sie im Detail.

Um eine spätere quantitative Analyse von Modellen und Messdaten zu ermöglichen, wurde die Theorie der Inversen Modellierung von dynamischen Modellen behandelt. Hierbei wurde besondere Aufmerksamkeit auf die asymptotische Theorie nichtlinearer Least Squares Schätzmethoden und auf dafür passende Optimierungsalgorithmen verwendet.

Ein einparametriges Transportmodell in Kombination mit einem Oberflächenreaktionsmodell, bestehend aus 5 Parametern, wird benutzt um das Verhalten eines Geassensors, bestehend aus einem Netzwerk von Nanodrähten, in einer Edelgasatmosphäre zu simulieren. Hierzu wurden die Reaktionskonstanten und Parameter beider Modelle durch einen nichtlinearen Least Sqares Schätzprozesses gefittet, unter zuhilfenahme von Inverser Modellierung.

Curriculum Vitae

Personal data

Name:	Marina Rehrl
Date of Birth:	9^{th} of March, 1985
Place of Birth:	Braunau am Inn, Upper Austria
Citizenship:	Austria

Education

1999 to 2004:	Federal Secondary College of Engineering, Linz Graphic and Communicational Design
July 2004:	Diploma thesis in the course of graduation with honor "Creation of a Corporate Identity for an oral surgeon"
since Autumn 2004:	Diploma Studies of Mathematics, University of Vienna Applied mathematics and scientific computing

Work Experience

July to August 2006:	Erste Bank Holding, Vienna Group Rating Methods
September 2009:	Allianz Elementar Versicherung, Vienna Realtor Service
March to June 2010 and September 2010 to March 2011:	Austrian Financial Market Authority, Vienna Actuarial Issues and Models